

Protein Sequence Searches - February 2005

All of the sequence databases on ABSS have recently been updated.

- Please note that the curators of the UniProt database have purged some temporary accession numbers from the most recent version of UniProt. These sequences have been assigned new permanent accession numbers. The new UniProt record may not contain the previous temporary accession number.
- If you encounter an accession number from an older search run against UniProt (results file extension **.rup**) that can no longer be found in the database, the permanent record with the new accession number can be found by searching the old accession number in the UniProt Protein, Archive database (UniPARC) at:

<http://www.pir.uniprot.org/database/archive.shtml>

If you have any questions regarding this information or your results, please contact any STIC searcher.

When submitting sequence search results for scanning into IFW, please include a copy of this attachment to assist any future Examiners or members of the public who may encounter UniProt temporary accession numbers.

This Page Blank (uspto)

GenCore version 5.1.6
Copyright (c) 1993 - 2005 CompuGen Ltd.

OM protein - protein search, using sw model

Run on: November 15, 2005, 12:07:58 ; Search time 126.737 Seconds
(without alignments)
48.827 Million cell updates/sec

Title: US-10-820-998-1

Perfect score: 78

Sequence: 1 SRLDYKSSLHLHGR 16

Scoring table: BLOSUM62

Gapop 10.0 , Gapext 0.5

Searched: 2105692 seqs, 386760381 residues

Total number of hits satisfying chosen parameters: 2105692

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database : A_Geneseq_16Dec04:*

1: Geneseqp1980s:*

2: Geneseqp1990s:*

3: Geneseqp2000s:*

4: Geneseqp2001s:*

5: Geneseqp2002s:*

6: Geneseqp2003as:*

7: Geneseqp2003bs:*

8: Geneseqp2004s:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	60	76.9	12	5	AU96886 Troponin
2	47	60.3	12	5	AU96888 Troponin
3	44	56.4	126	8	ABO54710 Human gen
4	44	56.4	126	8	ABO58262 Human gen
5	44	56.4	695	8	ABM83818 Human dia
6	44	56.4	698	8	ADQ39857 Human myo
7	44	56.4	725	8	ADQ39858 Human myo
8	42	53.8	505	7	ADB65266 Human pro
9	42	53.8	558	7	ADI21141 Novel hum
10	42	53.8	561	6	ABR58441 Human NOV
11	41	52.6	217	6	ABU16735 Protein e
12	41	52.6	823	4	ABB71741 Drosophil
13	41	52.6	902	8	ADS43963 Bacterial
14	41	52.6	960	6	ABR58216 Human nil
15	40	51.3	14	4	AAB87421 Human gen
16	40	51.3	20	5	AU956648 Lung tumo
17	40	51.3	20	5	AU956649 Lung tumo
18	40	51.3	20	6	ABU65523 Lung canc
19	40	51.3	20	6	ABU66524 Lung canc
20	40	51.3	20	7	ADH47480 Human lun
21	40	51.3	20	7	ADH47481 Human lun
22	40	51.3	20	7	ADH47670 Human lun
23	40	51.3	20	8	ADJ21589 Human lun
24	40	51.3	20	8	ADJ21399 Human lun
25	40	51.3	20	8	ADJ21400 Human lun

26	40	51.3	30	7	ADH47526 Human lun
27	40	51.3	30	8	ADJ21445 Human lun
28	40	51.3	47	5	AU95564 Lung tumo
29	40	51.3	47	6	ABU69536 Human lun
30	40	51.3	47	6	ABU66439 Lung canc
31	40	51.3	47	7	ADH47351 Human lun
32	40	51.3	47	8	ADJ21270 Human lun
33	40	51.3	81	5	AAM51134 Melanoma
34	40	51.3	96	5	AU95563 Lung tumo
35	40	51.3	96	6	ABU69535 Human lun
36	40	51.3	96	6	ABU66438 Lung canc
37	40	51.3	96	7	ADH47350 Human lun
38	40	51.3	96	8	ADJ21269 Human lun
39	40	51.3	108	4	AAB76886 Human lun
40	40	51.3	108	4	AAB76870 Human lun
41	40	51.3	108	5	AU95525 L552S lun
42	40	51.3	108	5	AU95521 L552S lun
43	40	51.3	108	6	ABU69497 Human lun
44	40	51.3	108	6	ABU69493 Human lun
45	40	51.3	108	6	ABU66395 Lung canc

ALIGNMENTS

RESULT 1
AAU96886
ID AAU96886 standard; peptide; 12 AA.

XX AC AAU96886;

XX DT 27-AUG-2002 (first entry)

XX DE Troponin C, TnC, derived peptide #2.

XX KW Troponin C; TnC; target chemical compound; peptide binding.

XX OS Unidentified.

XX PN JP2002058479-A.

XX PD 26-FEB-2002.

XX PF 14-AUG-2000; 2000JP-00245677.

XX PR 14-AUG-2000; 2000JP-00245677.

XX PA (CANO) CANON KK.

XX DR WPI; 2002-447068/48.

XX PT Determination and isolation of a structure recognizing amino acid sequence that is capable of recognition of a target chemical substance.
XX PS Example 2; Page 10; 18pp; Japanese.
XX CC The invention relates to the determination of a structure recognizing amino acid sequence useful as a peptide capable of recognition and selective binding with a target chemical compound in a living sample, comprising: (1) screening of a peptide fraction solely adsorbed on a carrier for the screening using 1st screening carrier with immobilised target chemical substance from variable random amino acid sequence region; (2) screening of the peptide fraction, excluding peptide fraction adsorbed on 2nd screening carrier from those selectively immobilised other than the target chemical substance in the sample, from the peptide groups adsorbed on the 1st screening step; (3) determination of the screened amino acid sequence in the 2nd step capable of binding with the target chemicals isolated in the 2nd step; and (4) determination of the aimed amino acid sequence capable of structure recognition in the elucidated peptides prepared by the preceding steps. The method is used for selective screening of a peptide capable of binding with the target chemical substance. The present sequence is a Troponin C (TnC) derived peptide used in an experiment demonstrating the method of the invention

```

XX SQ Sequence 12 AA;
Query Match 76.9%; Score 60; DB 5; Length 12;
Best Local Similarity 100.0%; Pred. No. 0.00099;
Matches 12; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 3 LDYKSSLLHLG 14
   |||||
Db 1 LDYKSSLLHLG 12

RESULT 2
AAU96888
ID AAU96888 standard; peptide; 12 AA.
XX
AC AAU96888;
XX
DT 27-AUG-2002 (first entry)
XX
DE Troponin C, TnC, derived peptide #4.
XX
KW Troponin C; TnC; target chemical compound; peptide binding.
XX
OS Unidentified.
XX
PN JP2002058479-A.
XX
PD 26-FEB-2002.
XX
PF 14-AUG-2000; 2000JP-00245677.
XX
PR 14-AUG-2000; 2000JP-00245677.
XX
PA (CANO ) CANON KK.
XX
DR WPI; 2002-447068/48.
XX
PT Determination and isolation of a structure recognizing amino acid
sequence that is capable of recognition of a target chemical substance.
XX
PS Example 2; Page 10; 18pp; Japanese.
XX
CC The invention relates to the determination of a structure recognising
amino acid sequence useful as a peptide capable of recognition and
selective binding with a target chemical compound in a living sample,
comprising: (1) screening of a peptide fraction solely adsorbed on a
carrier for the screening using 1st screening carrier with immobilised
target chemical substance from variable random amino acid sequence region
; (2) screening of the peptide fraction, excluding peptide fraction
adsorbed on 2nd screening carrier from those selectively immobilised
other than the target chemical substance in the sample, from the peptide
groups adsorbed on the 1st screening step; (3) determination of the
screened amino acid sequence in the 2nd step capable of binding with the
target chemicals isolated in the 2nd step; and (4) determination of the
aimed amino acid sequence capable of structure recognition in the
elucidated peptides prepared by the preceding steps. The method is used
for selective screening of a peptide capable of binding with the target
chemical substance. The present sequence is a Troponin C (TnC) derived
peptide used in an experiment demonstrating the method of the invention
XX
SQ Sequence 12 AA;
Query Match 60.3%; Score 47; DB 5; Length 12;
Best Local Similarity 81.8%; Pred. No. 0.21;
Matches 9; Conservative 2; Mismatches 0; Indels 0; Gaps 0;

Qy 3 LDYKSSLLHL 13
   |||||
Db 1 LDYKSSLLHL 11

RESULT 3

```

```

ABO54710
ID ABO54710 standard; protein; 126 AA.
XX
AC ABO54710;
XX
DT 29-JUL-2004 (first entry)
XX
DE Human genome derived single exon protein #944.
XX
KW Human; gene expression; single exon probe; microarray;
alternative splicing event; genomic alteration.
XX
OS Homo sapiens.
XX
PN US2003194704-A1.
XX
PD 16-OCT-2003.
XX
PF 03-APR-2002; 2002US-00029386.
XX
PR 03-APR-2002; 2002US-00029386.
XX
PA (PENN/) PENN S G.
PA (RANK/) RANK D R.
PA (HANZ/) HANZEL D K.
XX
PI Penn SG, Rank DR, Hanzel DK;
XX
DR WPI; 2004-119264/12.
XX
PT New human genome-derived single exon nucleic acid probes useful for human
gene expression analysis, for identifying or characterizing alternative
splicing events, for assessing genomic alterations or as tools for
surveying tissues.
XX
PS Claim 45; SEQ ID NO 28344; 80pp; English.
XX
CC The invention relates to a nucleic acid probe for measuring human gene
expression, comprising any of the 27,400 fully defined nucleotide
sequences in the specification, or their complements or fragments, and
encoding at least 8 amino acids of any of the 6888 amino acid sequences
fully defined in the specification. The probe is a single exon probe that
hybridises under high stringency conditions to a nucleic acid molecule
expressed in human cells or tissues. Also included are a spatially-
addressable set of single exon nucleic acid probes for measuring human
gene expression (comprising a plurality of single exon nucleic acid
probes cited above, where each of the plurality of probes is separately
and addressably isolatable or amplifiable from the plurality), a single
exon microarray for measuring human gene expression, a method of
measuring human gene expression, a vector comprising the single exon
probe cited above, an ORF-encoded peptide comprising at least 8
contiguous amino acids of any of the above-mentioned amino acid
sequences (optionally with conservative amino acid substitutions), an
isolated antibody that binds specifically to a peptide cited above,
methods of selling and/or licensing single exon probes or microarrays to
a customer desiring to measure gene expression, a method of providing
human gene expression data by subscription, and a computer-readable
storage medium which contains a database having a plurality of records
(each record including data on the expression of a single exon probe
cited above. The probe, methods and apparatus are useful in gene
expression analysis. The probes may be used as tools for surveying
tissues to detect the presence of expressed messages that contain their
specific exon, or in constructing genome-derived single exon microarrays.
In addition, the probes are used in identifying and characterising
alternations in the genomic locus that includes their exon, in assessing
smaller genomic alterations, in priming the synthesis of nucleic acids,
or in expressing the ORF-encoded peptide. The present sequence is a human
single exon probe protein of the invention. Note: The sequence data for
this patent did not form part of the printed specification, but was
obtained in electronic format directly from USPTO at
seqdata.uspto.gov/sequence.html?docID=20030194704
XX

```

CC	tissues to detect the presence of expressed messages that contain their	
CC	specific exon, or in constructing genome-derived single exon microarrays.	
CC	In addition, the probes are used in identifying and characterising	
CC	alternative splicing events, in detecting and characterising gross	
CC	alterations in the genomic locus that includes their exon, in assessing	
CC	smaller genomic alterations, in priming the synthesis of nucleic acids,	
CC	or in expressing the ORF-encoded peptide. The present sequence is a human	
CC	single exon probe protein of the invention. Note: The sequence data for	
CC	this patent did not form part of the printed specification, but was	
CC	obtained in electronic format directly from USPTO at	
CC	ccsqdata.uspto.gov/sequence.html?DocID=20030194704	
XX		
QQ	Sequence 126 AA;	
	Query Match 56.4%; Score 44; DB 8; Length 126;	
	Best Local Similarity 56.2%; Pred. No. 11;	
	Matches 9; Conservative 3; Mismatches 4; Indels 0; Gaps 0;	
QY	1 SRLDYLKSSLLHLGSR 16	
	: :	
DB	41 SRTDLVKSELLHIESQ 56	
RESULT 5		
ABM83818		
ID	ABM83818 standard; protein; 695 AA.	
XX		
AC	ABM83818;	
XX		
DT	18-NOV-2004 (first entry)	
XX		
DE	Human diagnostic and therapeutic pprotein SEQ ID NO:4067.	
XX		
KW	gene therapy; human diagnostic and therapeutic polynucleotide; dithp.	
XX		
OS	Homo sapiens.	
XX		
PN	WO2004023973-A2.	
XX		
PD	25-MAR-2004.	
XX		
PF	12-SEP-2003; 2003WO-US028227.	
XX		
PR	12-SEP-2002; 2002US-0410259P.	
XX		
PR	12-SEP-2002; 2002US-0410260P.	
XX		
PA	(INCY-) INCYTE CORP.	
XX		
PI	Schmidt JP, Wright RJ, Bruns CM, Marjanovic MM, Shen F;	
PI	Harthshorne TA, Suchorolski MT, Altus CM, Pitts SJ, Elder LV;	
PI	Mooney EM, Delegeane AM, Panegar IS, Banville SC, Reddy TP;	
PI	Stevens KA, Blanchard JL, Panzer SR, Wang X, Au AP, Gerstein EH;	
PI	Peralta CH, Anderson SB, Rioux P, Shen EJ, Wu MC, Stuve LL;	
PI	Lagace RE, Spiro PA, Stewart EA, Wingrove J, Vitt UA, Kirton ES;	
PI	Xu Y, Kwong M, Pollicky JL, Hurwitz BL, Ma Y, Jackson JL, Gietzen D;	
PI	Patry S, Shi X, Suarez CJ;	
XX		
DR	WPI: 2004-329368/30.	
DR	N-PSDB; ACN42470.	
XX		
PT	New diagnostic and therapeutic polynucleotides and polypeptides, useful	
PT	in diagnosing a condition, disease or disorder associated with human	
PT	molecules, e.g. autoimmune or inflammatory disorders, in gene therapy or	
PT	in gene mapping.	
XX		
PS	Claim 27; Page; 190pp; English.	
XX		
CC	The invention relates to novel diagnostic and therapeutic polynucleotides	
CC	selected from one of the 2722 sequences defined in the specification. A	
CC	polynucleotide of the invention may have a use in gene therapy. The human	
CC	diagnostic and therapeutic polynucleotides (dithp) or polypeptides may be	
CC	used to diagnose a particular condition, disease or disorder associated	
CC	with human molecules, e.g. cell proliferative disorders.	

CC autoimmune/inflammatory disorder, developmental disorder, endocrine
CC disorder, neurological disorders, gastrointestinal disorders, or
CC infections caused by virus, bacteria, fungi or parasite. The dithp
CC molecules may also be used in genetic mapping, in identifying individuals
CC from minute biological samples, in detecting single nucleotide
CC polymorphisms, as molecular weight markers, and for somatic or germline
CC gene therapy. The present sequence represents a dithp protein of the
CC invention. Note: The sequence data for this patent is not represented in
CC the printed specification, but was obtained in electronic format directly
CC from WIPO at www.wipo.int/pct/en/sequences/listing.htm
XX
SQ Sequence 695 AA;

Query Match 56.4%; Score 44; DB 8; Length 695;
Best Local Similarity 56.2%; Pred. No. 77;
Matches 9; Conservative 3; Mismatches 4; Indels 0; Gaps 0;

QY 1 SRLDYLKSSLLHLGSR 16
Db 202 SRTDLVKSELLHIESQ 217
||| :||| ||| :|

RESULT 6
ADQ39857
ID ADQ39857 standard; protein; 698 AA.
XX AC ADQ39857;
XX AC
XX AC
DT 18-NOV-2004 (first entry)
XX Human myocardial infarction-associated gene derived protein, SEQ ID 1520.
DE
XX Myocardial infarction; detection; single nucleotide polymorphism; SNP;
KW cardiant; gene therapy; human.
XX Homo sapiens.
XX WO2004058052-A2.
XX 15-JUL-2004.
XX 22-DEC-2003; 2003WO-US040978.
XX 20-DEC-2002; 2002US-0434778P.
PR 10-MAR-2003; 2003US-0453135P.
PR 30-APR-2003; 2003US-0466412P.
PR 23-SEP-2003; 2003US-0504955P.
XX (APPL-) APPLERA CORP.
XX Cargill M, Devlin JJ, Takoubova O;
XX WPI; 2004-533949/51.
DR N-PSDB; ADQ39029.
XX
XX Identifying an individual who has an altered risk for developing
PT myocardial infarction by detecting a single nucleotide polymorphism in
PT the individual's nucleic acids.
XX Claim 10; SEQ ID NO 1520; 145pp; English.
XX The invention relates to a novel method for identifying an individual who
CC has an altered risk for developing myocardial infarction. The method
CC comprises detecting a single nucleotide polymorphism (SNP) in any one of
CC the nucleotide sequences given in the specification in the individual's
CC nucleic acids, where the presence of the SNP is correlated with an
CC altered risk for myocardial infarction in the individual. The invention
CC further comprises: an isolated nucleic acid molecule comprising at least
CC 8 contiguous nucleotides where one of the nucleotides is an SNP given in
CC the specification or its complement and encoding any one of the amino
CC acid sequences given in the specification; an isolated polypeptide
CC comprising an amino acid sequence given in the specification; an antibody
CC that specifically binds to the polypeptide or its antigen-binding

CC fragment; an amplified polynucleotide containing an SNP given in the
CC specification and which is between about 16 and 1000 nucleotides in
CC length; a kit for detecting an SNP in a nucleic acid, comprising the
CC polynucleotide, a buffer and an enzyme; a method of detecting an SNP in a
CC nucleic acid molecule; a method of detecting a variant polypeptide; and a
CC method for identifying an agent useful in treating or preventing
CC myocardial infarction. The novel detection method has cardiant activity.
CC The nucleic acids of the invention may be used in gene therapy. The
CC method is useful in identifying an individual who has an increased or
CC decreased risk for developing myocardial infarction and for preparing a
CC composition for treating or preventing myocardial infarction. This
CC sequence represents the protein of a human myocardial infarction-
CC associated gene containing one or more SNP's of the invention. Note: This
CC sequence was not shown in the specification. The sequence has come from
CC an electronic sequence listing downloaded from the WIPO website.
XX
SQ Sequence 698 AA;

Query Match 56.4%; Score 44; DB 8; Length 698;
Best Local Similarity 56.2%; Pred. No. 78;
Matches 9; Conservative 3; Mismatches 4; Indels 0; Gaps 0;

QY 1 SRLDYLKSSLLHLGSR 16
Db 175 SRTDLVKSELLHIESQ 190
||| :||| ||| :|

RESULT 7
ADQ39858
ID ADQ39858 standard; protein; 725 AA.
XX AC ADQ39858;
XX AC
XX 18-NOV-2004 (first entry)
XX Human myocardial infarction-associated gene derived protein, SEQ ID 1521.
DE
XX Myocardial infarction; detection; single nucleotide polymorphism; SNP;
KW cardiant; gene therapy; human.
XX Homo sapiens.
XX WO2004058052-A2.
XX 15-JUL-2004.
XX 22-DEC-2003; 2003WO-US040978.
XX 20-DEC-2002; 2002US-0434778P.
PR 10-MAR-2003; 2003US-0453135P.
PR 30-APR-2003; 2003US-0466412P.
PR 23-SEP-2003; 2003US-0504955P.
XX (APPL-) APPLERA CORP.
XX Cargill M, Devlin JJ, Takoubova O;
XX WPI; 2004-533949/51.
DR N-PSDB; ADQ39030.
XX
XX Identifying an individual who has an altered risk for developing
PT myocardial infarction by detecting a single nucleotide polymorphism in
PT the individual's nucleic acids.
XX Claim 10; SEQ ID NO 1521; 145pp; English.
XX The invention relates to a novel method for identifying an individual who
CC has an altered risk for developing myocardial infarction. The method
CC comprises detecting a single nucleotide polymorphism (SNP) in any one of
CC the nucleotide sequences given in the specification in the individual's
CC nucleic acids, where the presence of the SNP is correlated with an
CC altered risk for myocardial infarction in the individual. The invention
CC further comprises: an isolated nucleic acid molecule comprising at least

CC 8 contiguous nucleotides where one of the nucleotides is an SNP given in
 CC the specification or its complement and encoding any one of the amino
 CC acid sequences given in the specification; an isolated polypeptide
 CC comprising an amino acid sequence given in the specification; an antibody
 CC that specifically binds to the polypeptide or its antigen-binding
 CC fragment; an amplified polynucleotide containing an SNP given in the
 CC specification and which is between about 16 and 1000 nucleotides in
 CC length; a kit for detecting an SNP in a nucleic acid, comprising the
 CC polynucleotide, a buffer and an enzyme; a method of detecting an SNP in a
 CC nucleic acid molecule; a method of detecting a variant polypeptide; and a
 CC method for identifying an agent useful in treating or preventing
 CC myocardial infarction. The novel detection method has cardiant activity.
 CC The nucleic acids of the invention may be used in gene therapy. The
 CC method is useful in identifying an individual who has an increased or
 CC decreased risk for developing myocardial infarction and for preparing a
 CC composition for treating or preventing myocardial infarction. This
 CC sequence represents the protein of a human myocardial infarction-
 CC associated gene containing one or more SNPs of the invention. Note: This
 CC sequence was not shown in the specification. The sequence has come from
 CC an electronic sequence listing downloaded from the WIPO website.

XX SQ Sequence 725 AA;

Query Match 56.4%; Score 44; DB 8; Length 725;
 Best Local Similarity 56.2%; Pred. No. 81; Mismatches 4; Indels 0; Gaps 0;
 Matches 9; Conservative 3; Mismatches 4; Indels 0; Gaps 0;

QY 1 SRLDYKSSLLHLGSR 16
 DB 202 SRTDLVKSELHIESQ 217
 |||:|||||:

RESULT 8
 ADB65266
 ID ADB65266 standard; protein; 505 AA.

XX AC ADB65266;
 XX DT 04-DEC-2003 (first entry)
 XX DE Human protein encoded by clone TESTI20055880.

XX Human; pharmaceutical; diagnostic; gene therapy; tissue regeneration;
 KW cell regeneration; membrane protein; signal transduction-related protein;
 KW transcription-related protein; osteoporosis; neurological disease;
 KW cancer; tumour.

XX OS Homo sapiens.

XX EP1308459-A2.

XX PD 07-MAY-2003.

XX PF 28-MAR-2002; 2002EP-00007401.

XX PR 05-NOV-2001; 2001JP-00379298.

XX PR 25-JAN-2002; 2002US-00350978.

XX (HELI-) HELIX RES INST.

PA (REAS-) RES ASSOC BIOTECHNOLOGY.

XX Isogai T, Sugiyama T, Otsuki T, Wakamatsu A, Sato H, Ishii S;
 PI Yamamoto J, Isono Y, Hio Y, Otsuka K, Nagai K, Irie R, Tamechika I;
 PI Seki N, Yoshikawa T, Otsuka M, Nagahari K, Masuho Y;

XX WPI; 2003-450961/43.

DR N-PSDB; ADB63296.

XX New polynucleotides and polypeptides, useful for developing a diagnostic
 PT marker or medicines for regulation of their expression and activity, or
 PT as targets of gene therapy.

XX Claim 1; Page; 222pp; English.

XX The invention discloses a polynucleotide comprising a sequence selected
 CC from 1970 fully defined nucleotide sequences which encode novel
 CC polypeptides. Also claimed is a polypeptide encoded by the polynucleotide
 CC or its partial peptide, an antibody binding to the polypeptide or peptide
 CC of the polynucleotide, immunologically assaying the polypeptide or
 CC peptide of the polynucleotide by contacting the polypeptide or peptide
 CC with the antibody of the encoded protein, and observing the binding
 CC between the two, a transformant carrying the polynucleotide in an
 CC expressible manner and an antisense polynucleotide. The oligonucleotide
 CC is useful as a primer for synthesising the polynucleotide, or as a probe
 CC for detecting the polynucleotide. The polynucleotides and encoded
 CC proteins are useful as pharmaceutical agents and many disease-related
 CC medicines may be included in them, for developing a diagnostic marker or
 CC medicines for regulation of their expression and activity, or as targets
 CC of gene therapy. The genes are involved in tissue and/or cell
 CC regeneration. Membrane proteins, signal transduction-related proteins,
 CC transcription-related proteins, disease-related proteins and genes
 CC encoding them can be used as indicators for diseases (e.g. osteoporosis,
 CC neurological diseases, cancer, tumours. The cDNA may be used to regulate
 CC the activity or expression of the encoded protein to treat diseases. The
 CC sequence presented is a protein of the invention. Note: Some of the
 CC sequence data for this patent is not represented in the printed
 CC specification, but is based on sequence information supplied by the
 CC European Patent Office.

XX SQ Sequence 505 AA;

Query Match 53.8%; Score 42; DB 7; Length 505;
 Best Local Similarity 64.3%; Pred. No. 1.2e+02;
 Matches 9; Conservative 2; Mismatches 3; Indels 0; Gaps 0;

QY 1 SRLDYKSSLLHLG 14
 DB 16 SRLDFLKASHFSLG 29
 |||:|||||:

RESULT 9
 ADI21141

ID ADI21141 standard; protein; 558 AA.

XX AC ADI21141;

XX DT 15-APR-2004 (first entry)

XX DE Novel human protein #116.

XX KW forensic; nutritional source; damaged tissue; diseased tissue;
 KW myeloid cell disorder; lymphoid cell disorder;

KW bone cartilage tissue growth; tendon tissue growth;

KW ligament tissue growth; nerve tissue growth; regeneration; wound healing;

KW tissue repair; tissue replacement; burn; incision; ulcer; cancer; human.

XX OS Homo sapiens.

XX PN WO2003025148-A2.

XX PD 27-MAR-2003.

XX PF 19-SEP-2002; 2002WO-US029964.

XX PR 19-SEP-2001; 2001US-0323739P.

XX PR 13-SEP-2002; 2002US-00323739.

XX (HYSE-) HYSEQ INC.

XX Tang YT, Asundi V, Goodrich RW, Ren F, Zhang J, Zhao QA, Wang J;
 PI Ghosh M, Xue AJ, Wehrman T, Weng G, Zhou P, Drmanac RT, Wang D;
 PI Haley-Vicente D;

XX WPI; 2003-354603/33.

DR N-PSDB; ADI21857.

XX

PT New polynucleotides and secreted proteins, useful for treating myeloid or
 PT lymphoid cell disorders, in bone cartilage, tendon, ligament and nerve
 PT tissue growth or regeneration, in wound healing, and in tissue repair and
 PT replacement.

XX Claim 20; SEQ ID NO 392; 156pp; English.

CC The invention relates to an isolated polynucleotide encoding a
 CC polypeptide with biological activity. The polynucleotides and
 CC polypeptides are useful in diagnostics, forensics, gene mapping,
 CC identification of mutations responsible for genetic disorders and other
 CC traits, to assess biodiversity, as nutritional sources or supplements.
 CC The polynucleotides may also be used as molecular weight markers,
 CC chromosome markers or map related gene positions, or as an antigen to
 CC raise anti-DNA antibodies or elicit immune response. The polypeptides are
 CC useful for raising antibodies, as markers for tissues in which the
 CC corresponding polypeptide is expressed, for re-engineering damaged or
 CC diseased tissues, for treating myeloid or lymphoid cell disorders, in
 CC bone cartilage, tendon, ligament and/or nerve tissue growth or
 CC regeneration, in wound healing, in tissue repair and replacement, in
 CC healing of burns, incisions and ulcers, and in treating cancer. The
 CC present sequence represents the amino acid sequence of a novel human
 CC protein.

XX Sequence 558 AA;

Query Match 53.8%; Score 42; DB 7; Length 558;

Best Local Similarity 53.3%; Pred. No. 1.4e+02; Indels 0; Gaps 0;

Matches 8; Conservative 2; Mismatches 5;

QY 2 RLDYKSSLLHLGSR 16

Db 528 RCDYLNKLSHIKQR 542

RESULT 10

ABR58441
 ID ABR58441 standard; protein; 561 AA.

XX ABR58441;

XX 07-JUL-2003 (first entry)

XX Human NOV43a.

XX Human; NOV; antidiabetic; anorectic; antibacterial; virucide;
 KW immunomodulator; cyostatic; nootropic; neuroprotective; dyslipidaemia;
 KW antiparkinsonian; antilipaemic; gene therapy; metabolic disorder;
 KW diabetes; obesity; infection; cachexia; cancer; Parkinson's disease;
 KW neurodegenerative disorder; Alzheimer's disease; immune disorder;
 KW haematopoietic disorder.

XX Homo sapiens.

XX WO2003029423-A2.

XX 10-APR-2003.

XX 02-OCT-2002; 2002WO-US031358.

XX 02-OCT-2001; 2001US-0326483P.

XX 05-OCT-2001; 2001US-0327342P.

XX 09-OCT-2001; 2001US-0327917P.

XX 09-OCT-2001; 2001US-0328029P.

XX 09-OCT-2001; 2001US-0328044P.

XX 09-OCT-2001; 2001US-0328056P.

XX 12-OCT-2001; 2001US-0328849P.

XX 15-OCT-2001; 2001US-0329414P.

XX 17-OCT-2001; 2001US-0330142P.

XX 22-OCT-2001; 2001US-0341058P.

PR 01-NOV-2001; 2001US-0346357P.

PR 12-APR-2002; 2002US-0371972P.

PR 12-APR-2002; 2002US-0371980P.

PR 17-APR-2002; 2002US-0373261P.

PR 19-APR-2002; 2002US-0373805P.

PR 23-APR-2002; 2002US-0374738P.

PR 16-MAY-2002; 2002US-0381101P.

PR 17-MAY-2002; 2002US-0381635P.

PR 20-MAY-2002; 2002US-0383830P.

PR 01-OCT-2002; 2002US-00262839.

XX (CURA-) CURAGEN CORP.

XX

PI Alsbrook JP, Anderson DW, Boldog FL, Burgess CE, Catterton E;

PI Edinger SR, Ellerman K, Gerlach VL, Gorman L, Guo X, Ji W;

PI Kekuda R, Leach MD, Li L, Miller CE, Patturajan M, Rieger DK;

PI Rothenberg ME, Shimkets RA, Smithson G, Spytek KA, Taupier RJ;

PI Vernet CAM, Voss EZ, Zerhusen BD, Zhong M;

XX WPI; 2003-381625/36.

DR N-PSDB; ACC72153.

XX

PT NOVX polypeptides and nucleic acids useful for diagnosing, preventing or
 PT treating NOVX-associated disorders, e.g. diabetes, obesity, cancer or
 PT dyslipidemia, and in chromosome mapping, tissue typing or
 PT pharmacogenomics.

XX Claim 1; Page 240; 487pp; English.

XX The present invention relates to novel human NOV proteins and their

CC coding sequences (ACC72075-ACC72181 and ABR58363-ABR58469). The NOV
 CC proteins are useful in manufacturing a medicament for treating a syndrome
 CC associated with a human disease. The NOV proteins and coding sequences
 CC may be used to diagnose, treat or prevent metabolic disorders such as
 CC diabetes or obesity, infections, cachexia, cancer, neurodegenerative
 CC disorders such as Alzheimer's disease or Parkinson's disease, immune
 CC disorders, haematopoietic disorders and various dyslipidaemias

XX Sequence 561 AA;

Query Match 53.8%; Score 42; DB 6; Length 561;

Best Local Similarity 53.3%; Pred. No. 1.4e+02;

Matches 8; Conservative 2; Mismatches 5; Indels 0; Gaps 0;

QY 2 RLDYKSSLLHLGSR 16

Db 531 RCDYLNKLSHIKQR 545

RESULT 11

ABU16735

ID ABU16735 standard; protein; 217 AA.

XX ABU16735;

XX 19-JUN-2003 (first entry)

XX

DE Protein encoded by Prokaryotic essential gene #2262.

XX Antisense; prokaryotic essential gene; cell proliferation; drug design.

XX Acinetobacter baumannii.

XX WO20027183-A2.

XX

XX 03-OCT-2002.

XX

XX 21-MAR-2002; 2002WO-US009107.

XX

XX 21-MAR-2001; 2001US-00815242.

PR 06-SEP-2001; 2001US-00948993.

PR 25-OCT-2001; 2001US-0342923P.

PR 08-FEB-2002; 2002US-00072851.

PR 06-MAR-2002; 2002US-0362699P.
 XX (ELIT-) ELITRA PHARM INC.
 XX Wang L, Zamudio C, Malone C, Haselbeck R, Ohlsen KL, Zyskind JW;
 PI Wall D, Trawick JD, Carr GJ, Yamamoto R, Forsyth RA, Xu HH;
 XX WPI; 2003-029926/02.
 DR N-PSDB; ACA20605.
 XX New antisense nucleic acids, useful for identifying proteins or screening
 PT for homologous nucleic acids required for cellular proliferation to
 PT isolate candidate molecules for rational drug discovery programs.
 XX
 XX Claim 25; SEQ ID NO 44659; 1766pp; English.
 PS The invention relates to an isolated nucleic acid comprising any one of
 XX the 6213 antisense sequences given in the specification where expression
 CC of the nucleic acid inhibits proliferation of a cell. Also included are:
 CC (1) a vector comprising a promoter operably linked to the nucleic acid
 CC encoding a polypeptide whose expression is inhibited by the antisense
 CC nucleic acid; (2) a host cell containing the vector; (3) an isolated
 CC polypeptide or its fragment whose expression is inhibited by the
 CC antisense nucleic acid; (4) an antibody capable of specifically binding
 CC the polypeptide; (5) producing the polypeptide; (6) inhibiting cellular
 CC proliferation or the activity of a gene in an operon required for
 CC proliferation; (7) identifying a compound that influences the activity of
 CC the gene product or that has an activity against a biological pathway
 CC required for proliferation, or that inhibits cellular proliferation; (8)
 CC identifying a gene required for cellular proliferation or the biological
 CC pathway in which a proliferation-required gene or its gene product lies
 CC or a gene on which the test compound that inhibits proliferation of an
 CC organism's activity; (9) manufacturing an antibiotic; (10) profiling a
 CC product is overexpressed or underexpressed; (12) determining the extent
 CC to which each of the strains is present in a culture or collection of
 CC strains; or (13) identifying the target of a compound that inhibits the
 CC proliferation of an organism. The antisense nucleic acids are useful for
 CC identifying proteins or screening for homologous nucleic acids required
 CC for cellular proliferation to isolate candidate molecules for rational
 CC drug discovery programs, or for screening homologous nucleic acids
 CC required for proliferation in cells other than *S. aureus*, *S. typhimurium*,
 CC *K. pneumoniae* or *P. aeruginosa*. The present sequence is encoded by one of
 CC the target prokaryotic essential genes. Note: The sequence data for this
 CC patent did not form part of the printed specification, but was obtained
 CC in electronic format directly from WIPO at
 CC ftp.wipo.int/pub/published_pct_sequences
 XX
 SQ Sequence 217 AA;
 Query Match 52.6%; Score 41; DB 6; Length 217;
 Best Local Similarity 69.2%; Pred. No. 69;
 Matches 9; Conservative 2; Mismatches 2; Indels 0; Gaps 0;
 QY 3 LDYKSSLLHLGS 15
 Db 197 LDYKANNALGS 209
 |||||:| |||
 RESULT 12
 ABB71741
 ID ABB71741 standard; protein; 823 AA.
 XX ABB71741;
 AC ABB71741;
 XX 26-MAR-2002 (first entry)
 DT Drosophila melanogaster polypeptide SEQ ID NO 42015.
 DE Drosophila melanogaster polypeptide; cell signalling; insecticide;
 KW Drosophila; developmental biology; cell signalling; insecticide;
 KW pharmaceutical.
 XX Drosophila melanogaster.
 OS

XX WO200171042-A2.
 XX 27-SEP-2001.
 XX 23-MAR-2001; 2001WO-US009231.
 XX 23-MAR-2000; 2000US-0191637P.
 PR 11-JUL-2000; 2000US-00614150.
 XX (PEKE) PE CORP NY.
 PA Venter JC, Adams M, Li PWD, Myers EW;
 PI WPI; 2001-656860/75.
 XX N-PSDB; ABL15844.
 DR New isolated nucleic acid detection reagent for detecting 1000 or more
 PT genes from *Drosophila* and for elucidating cell signaling and cell-cell
 PT interactions.
 XX Disclosure; SEQ ID NO 42015; 21pp + Sequence Listing; English.
 PS The invention relates to an isolated nucleic acid detection reagent
 CC capable of detecting 1000 or more genes from *Drosophila*. The invention is
 CC useful in developmental biology and in elucidating cell signalling and
 CC cell-cell interactions in higher eukaryotes for the development of
 CC insecticides, therapeutics and pharmaceutical drugs. The invention
 CC discloses genomic DNA sequences (ABL16176-ABL30511), expressed DNA
 CC sequences (ABL01840-ABL16175) and the encoded proteins (ABBS57737-
 CC ABB72072). The sequence data for this patent did not form part of the
 CC printed specification, but was obtained in electronic format directly
 CC from WIPO at ftp.wipo.int/pub/published_pct_sequences
 XX
 SQ Sequence 823 AA;
 Query Match 52.6%; Score 41; DB 4; Length 823;
 Best Local Similarity 75.0%; Pred. No. 3.2e+02;
 Matches 9; Conservative 0; Mismatches 3; Indels 0; Gaps 0;
 QY 2 RLDYKSSLLHL 13
 Db 85 RLDYKQHLLDL 96
 |||||:| |||
 RESULT 13
 ADS43963
 ID ADS43963 standard; protein; 902 AA.
 XX ADS43963;
 AC 02-DEC-2004 (first entry)
 DT Bacterial polypeptide #22393.
 XX Recombinant DNA construct; transformed plant; improved plant property;
 KW cold tolerance; heat tolerance; drought tolerance; herbicide; osmosis;
 KW pathogen tolerance; pest tolerance; plant disease resistance;
 KW cell cycle pathway modification; plant growth regulator;
 KW homologous recombination; seed oil yield; protein yield; carbohydrate;
 KW nitrogen; phosphorus; photosynthesis; lignin; galactomannan;
 KW bacterial polypeptide.
 XX Bacteria.
 OS US2003233675-A1.
 XX 18-DEC-2003.
 PD 20-FEB-2003; 2003US-00369493.
 PF 21-FEB-2002; 2002US-0360039P.
 XX

PA (CAOY/) CAO Y.
PA (HINK/) HINKLE G J.
PA (SLAT/) SLATER S C.
PA (CHEN/) CHEN X.
PA (GOLD/) GOLDMAN B S.
XX
PI Cao Y, Hinkle GJ, Slater SC, Chen X, Goldman BS;
XX WPI; 2004-061375/06.
XX
XX New recombinant DNA construct comprising a promoter positioned to provide
PT for expression of a polynucleotide encoding a polypeptide from a
PT microbial source, useful for producing plants with improved properties.
XX
XX Claim 1; SEQ ID NO 22393; 122pp; English.
XX
XX The invention relates to a recombinant DNA construct comprising a
CC promoter functional in a plant cell, where the promoter is positioned to
CC provide for expression of a polynucleotide encoding a polypeptide from a
CC microbial source. The invention also relates to a transformed plant
CC comprising the recombinant DNA construct and a method of producing a
CC transformed plant having an improved property. The plant is a crop plant
CC such as maize or soybean. The method of producing a transformed plant
CC having an improved property comprises transforming a plant with the
CC recombinant DNA construct and growing the transformed plant, where the
CC polynucleotide or polypeptide is useful for improving plant properties.
CC The recombinant DNA construct is useful for producing plants with
CC improved plant properties, e.g. improved cold, heat or drought tolerance,
CC tolerance to herbicides, extreme osmotic conditions, pathogens or pests,
CC increased resistance to plant disease, better growth rate by modification
CC of the cell cycle pathway with plant growth regulators, increased rate of
CC homologous recombination, modified seed oil or protein yield and/or
CC content, improved yield by modification of carbohydrate, nitrogen or
CC phosphorus use and/or uptake, by modification of photosynthesis or by
CC providing improved plant growth and development under at least one stress
CC condition. Improved lignin production or improved galactomannan
CC production. This sequence represents a bacterial polypeptide used in the
CC scope of the invention. Note: The sequence data for this patent did not
CC form part of the printed specification but was obtained in electronic
CC format from USPTO at seqdata.uspto.gov/sequence.html.
XX
XX Sequence 902 AA;
SQ
Query Match 52.6%; Score 41; DB 8; Length 902;
Best Local Similarity 66.7%; Pred. No. 3.6e+02;
Matches 8; Conservative 2; Mismatches 2; Indels 0; Gaps 0;
QY 5 YLKSSLLHLGSR 16
|:|:|:|:|:|
Db 48 YIKSPLVHLHSR 59
RESULT 14
ABR58216
ID ABR58216 standard; protein; 960 AA.
XX
XX ABR58216;
AC
XX 11-AUG-2003 (first entry)
DT
XX Human nil per os protein.
DE
XX Antinflammatory; Gastrointestinal; Vasotropic; Hepatotrophic; Antiulcer;
KW Cystostatic; nil per os protein; digestive system; human.
XX
XX Homo sapiens.
OS
XX WO2003007800-A2.
PN
XX 30-JAN-2003.
PD
XX 17-JUL-2002; 2002WO-US022904.
PF
XX

PR 17-JUL-2001; 2001US-0306319P.
XX (GEHO) GEN HOSPITAL CORP.
PA
XX Fishman MC, Mayer AN;
PI
XX WPI; 2003-278311/27.
DR N-PSDB; ACC80945.
DR
XX Determining wether a subject has, or is at risk, or developing a disease
PT related to a nil per os protein e.g. digestive system disorder or cancer,
PT comprises detecting mutations in the nil per os gene.
XX
XX Claim 16; Page 63-65; 65pp; English.
XX
XX The present invention relates to determining whether a test subject has,
CC or is at risk of developing, a disease or condition related to a nil per
CC os protein, comprising analyzing a nucleic acid to determine whether the
CC subject has a mutation in the gene for the nil per os protein, where the
CC presence of a mutation indicates that the subject has, or is at risk of
CC developing such a disease. The method is considered antinflammatory,
CC gastrointestinal, vasotropic, hepatotropic, antiulcer and cystostatic.
CC The method is useful for determining whether a test subject, especially a
CC human, has, or is at risk of developing, a disease or condition related
CC to a nil per os protein, particularly diseases or conditions of the
CC digestive system is useful for identifying a compound that can be used to
CC treat or prevent one of these diseases or conditions. The method is
CC useful for treating one of these diseases, and especially digestive organ
CC failure, where the patient has a mutation in a gene encoding a nil per os
CC protein. The methods can also be used to diagnose or treat the above
CC diseases in any mammal e.g. domestic pets and livestock. The present
CC sequence represents the nil per os protein of human
XX
XX Sequence 960 AA;
SQ
Query Match 52.6%; Score 41; DB 6; Length 960;
Best Local Similarity 53.3%; Pred. No. 3.9e+02;
Matches 8; Conservative 3; Mismatches 4; Indels 0; Gaps 0;
QY 1 SRLDYLKSSLLHLGS 15
|:|:|:|:|:|
Db 206 SMDYLKSKXVKGAS 220
RESULT 15
AAB87421
ID AAB87421 standard; peptide; 14 AA.
XX
XX AAB87421;
AC
XX 22-MAY-2001 (first entry)
DT
XX Human gene 3 encoded secreted protein fragment, SEQ ID NO:162.
DE
XX Human; secreted protein; proliferative disorder; cancer; tumour;
KW foetal abnormality; developmental abnormality; haematopoietic disorder;
KW immune system disorder; AIDS; autoimmune disease; rheumatoid arthritis;
KW inflammation; allergy; neurological disorder; Alzheimer's disease;
KW Parkinson's disease; cognitive disorder; schizophrenia; asthma;
KW skin disorder; psoriasis; sepsis; diabetes; atherosclerosis;
KW cardiovascular disorder; angioenic disorder; kidney disorder;
KW gastrointestinal disorder; pregnancy-related disorder;
KW endocrine disorder; infection; wound healing; vulnery; cell culture;
KW chemotaxis; food additive; binding partner identification.
XX
XX Homo sapiens.
OS
XX WO200118022-A1.
PN
XX 15-MAR-2001.
PD
XX 31-AUG-2000; 2000WO-US024008.
PF
XX

PR 03-SEP-1999; 99US-015231SP.
PR 03-SEP-1999; 99US-0152317P.
XX
XX (HUMA-) HUMAN GENOME SCI INC.
XX
XX Ni J, Baker KP, Birse CE, Fiscella M, Komatsoulis GA, Rosen CA;
PI Soppet DR, Young PE, Ebner R, Duan DR, Olsen HS, Lafleur DW;
PI Moore PA, Shi Y, Wei Y, Florence KA;
XX
XX WPI; 2001-203081/20.
DR
DR
XX
XX Nucleic acid molecules encoding human secreted proteins, used in
PT preventing, treating or ameliorating a disorder, e.g. Alzheimer's and
PT Parkinson's diseases and cancers.
XX
XX Disclosure; Page 18; 607pp; English.
PS
XX
XX AAF91858-AAF91929 represent cDNAs corresponding to 52 human secreted
CC protein genes, and AAB87342-AAB87413 represent the proteins they encode.
CC AAB87414-AAB87454 represent human secreted protein fragments. The genes
CC and their corresponding secreted proteins are useful for preventing,
CC treating or ameliorating medical conditions, e.g., by protein or gene
CC therapy. Pathological conditions can be diagnosed by determining the
CC amount of the new protein in a sample or by determining the presence of
CC mutations in the new genes. Specific uses are described for each of the
CC 52 genes, based on the tissues in which they are most highly expressed,
CC and include developing products for the diagnosis or treatment of
CC proliferative disorders, cancer, tumours, foetal and developmental
CC abnormalities, haematopoietic disorders, diseases of the immune system,
CC AIDS, autoimmune diseases (e.g., rheumatoid arthritis), inflammation,
CC allergies, neurological disorders (e.g., Alzheimer's disease,
CC Parkinson's disease), cognitive disorders, schizophrenia, asthma, skin
CC disorders (e.g., psoriasis), sepsis, diabetes, atherosclerosis,
CC cardiovascular disorders, angiogenic disorders, kidney disorders,
CC gastrointestinal disorders, pregnancy-related disorders, endocrine
CC disorders, and infections. The proteins can also be used to aid wound
CC healing and epithelial cell proliferation, to prevent skin aging due to
CC sunburn, to maintain organs before transplantation, for supporting cell
CC culture of primary tissues, to regenerate tissues, to identify their
CC cognate ligands or binding partners, and in chemotaxis, and can be used
CC as a food additive or preservative to modify storage properties.
CC Antibodies specific for a protein of the invention can be used in
CC alleviating symptoms associated with the disorders mentioned above, and
CC in diagnostic immunoassays e.g., radioimmunoassay or enzyme linked
CC immunosorbent assay (ELISA). The present sequence represents a human
CC secreted protein fragment referred to in the disclosure of the invention
XX
XX Sequence 14 AA;
SQ
Query Match 51.3%; Score 40; DB 4; Length 14;
Best Local Similarity 63.6%; Pred. No. 4.4;
Matches 7; Conservative 4; Mismatches 0; Indels 0; Gaps 0;
QY 6 LKSSLLHLGSR 16
|::|::|::|
Db 3 LRASLVHVGSR 13

Search completed: November 15, 2005, 12:22:51
Job time : 128.737 secs

This Page Blank (uspto)

GenCore version 5.1.6
Copyright (c) 1993 - 2005 Compugen Ltd.

OM protein - protein search, using sw model

Run on: November 15, 2005, 12:11:54 ; Search time 14.3158 Seconds
(without alignments)
107.536 Million cell updates/sec

Title: US-10-820-998-1
Perfect score: 78
Sequence: 1 SRLDYKSLHLGSR 16
Scoring table: BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 283416 seqs, 96216763 residues
Total number of hits satisfying chosen parameters: 283416
Minimum DB seq length: 0
Maximum DB seq length: 2000000000
Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : PIR_79:.*
1: pir1:.*
2: pir2:.*
3: pir3:.*
4: pir4:.*

Pred. No. is the number of results predicted by chance to have a
score greater than or equal to the score of the result being printed,
and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	48	61.5	442	2 T20638	hypothetical prote
2	44	56.4	847	2 C87678	TonB-dependent rec
3	43	55.1	216	2 F84023	transcription regu
4	43	55.1	986	2 S46812	hypothetical prote
5	41	52.6	403	2 C87315	hypothetical prote
6	41	52.6	507	2 B64433	probable O-antigen
7	41	52.6	565	2 D72222	conserved hypothet
8	41	52.6	648	2 A71647	glycine-tRNA ligas
9	41	52.6	772	2 H84605	hypothetical prote
10	41	52.6	902	2 S33318	dynammin-like prote
11	41	52.6	960	2 T17297	hypothetical prote
12	40.5	51.9	430	2 AF3460	histidinol dehydro
13	40	51.3	354	2 D83980	hypothetical prote
14	40	51.3	495	2 D86442	probable amino aci
15	40	51.3	601	2 T37738	hypothetical prote
16	39	50.0	288	2 E90064	conserved hypothet
17	39	50.0	531	2 T36377	probable exporter
18	39	50.0	1034	2 G90591	hypothetical prote
19	39	50.0	1369	2 T43433	alpha-glucan synth
20	39	50.0	1726	2 A39401	merozoite surface
21	38	48.7	109	2 C72459	hypothetical prote
22	38	48.7	139	2 I38620	zinc finger protei
23	38	48.7	174	2 A85979	hypothetical prote
24	38	48.7	174	2 A65106	hypothetical 19.7
25	38	48.7	174	2 F91133	hypothetical prote
26	38	48.7	238	2 H18959	probable periplasm
27	38	48.7	384	2 A10556	conserved hypothet
28	38	48.7	467	2 D88710	protein C43G2.1 [1
29	38	48.7	501	2 A55749	spliceosome-associ

30	38	48.7	531	2 T43551	multidrug resistan
31	38	48.7	683	2 JE0262	long-chain-fatty-a
32	38	48.7	699	2 B97049	anaerobic ribonuc
33	38	48.7	702	2 A96959	oxygen-sensitive r
34	38	48.7	988	2 T45717	receptor-kinase li
35	38	48.7	1174	2 T43051	protein kinase C (
36	38	48.7	1217	2 T39427	probable myosin I
37	38	48.7	3013	2 AB0480	probable myosin Y
38	37	47.4	124	1 MN1HMS	nonstructural prot
39	37	47.4	139	2 F97431	hypothetical prote
40	37	47.4	166	2 D37844	baIE 19.5K protein
41	37	47.4	302	1 A90448	phosphoesterase-re
42	37	47.4	394	2 T31891	hypothetical prote
43	37	47.4	404	2 T20453	hypothetical prote
44	37	47.4	404	2 I59589	starvation-sensing
45	37	47.4	413	2 T04520	hypothetical prote

ALIGNMENTS

RESULT 1

T20638

hypothetical protein T06H11.4 - Caenorhabditis elegans

C;Species: Caenorhabditis elegans

C;Date: 15-Oct-1999 #sequence_revision 15-Oct-1999 #text_change 12-Jul-2004

C;Accession: T20638; T24630

R;Kershaw, J.

Submitted to the EMBL Data Library, June 1995

A;Reference number: Z19303

A;Accession: T20638

A;Status: preliminary; translated from GB/EMBL/DDBJ

A;Molecule type: DNA

A;Residues: 1-442 <W1>

A;Cross-references: UNIPROT:Q19242; EMBL:Z49887; PIDN:CAA90060.1; GSPDB:GN00028; CESP:T06H11.4

A;Experimental source: Clone F0989

R;Kershaw, J.

Submitted to the EMBL Data Library, June 1995

A;Reference number: Z19914

A;Accession: T24630

A;Status: preliminary; translated from GB/EMBL/DDBJ

A;Molecule type: DNA

A;Residues: 1-442 <W1>

A;Cross-references: EMBL:Z49889; PIDN:CAA90069.1; GSPDB:GN00028; CESP:T06H11.4

A;Experimental source: clone T06H11

C;Genetics:

A;Gene: CESP.T06H11.4

A;Map position: X

A;Introns: 45/1; 95/3; 150/2; 208/3; 250/2; 292/3

C;Superfamily: molybdenum cofactor biosynthesis protein, MoeA type

Query Match 61.5%; Score 48; DB 2; Length 442;
Best Local Similarity 72.7%; Pred. No. 1.5;
Matches 8; Conservative 3; Mismatches 0; Indels 0; Gaps 0;

QY 4 DYLSKSLHLG 14

|||||

Db 280 DYMKALLHLG 290

RESULT 2

C87678

TonB-dependent receptor [imported] - Caulobacter crescentus

C;Species: Caulobacter crescentus

C;Date: 20-Apr-2001 #sequence_revision 20-Apr-2001 #text_change 09-Jul-2004

C;Accession: C87678

R;Nierman, W.C.; Feldblyum, T.V.; Paulsen, I.T.; Nelson, K.E.; Eisen, J.; Heidelberg, J.F.

B.; Laub, M.T.; DeBoy, R.T.; Dodson, R.J.; Durkin, A.S.; Gwinn, M.L.; Haft, D.H.; Kolon

n, J.; Ermolaeva, M.; White, O.; Salzberg, S.L.; Shapiro, L.; Venter, J.C.; Fraser, C.M.

Proc. Natl. Acad. Sci. U.S.A. 98, 4136-4141, 2001

A;Title: Complete Genome Sequence of Caulobacter crescentus.

A;Reference number: A87249; MUID:21173698; PMID:11259647

A;Accession: C87678

```

A:Status: preliminary
A:Molecule type: DNA
A:Residues: 1-847 <STO>
A:Cross-references: UNIPROT:Q9A2US; GB:AE005673; NID:gl3425181; PIDN:AAK25423.1; GSPDB:G
C:Genetics:
A:Gene: CC3461

Query Match      56.4%; Score 44; DB 2; Length 847;
Best Local Similarity 66.7%; Pred. No. 15;
Matches 8; Conservative 3; Mismatches 1; Indels 0; Gaps 0;

Qy 5 YLKSSLLHLGSR 16
   ||||| | | | |
Db 737 YLKASLQHVGNR 748

RESULT 3
F84023
transcription regulator involved in phosphate transport system BH2990 [imported] - Bacil
C:Species: Bacillus halodurans
C:Date: 01-Dec-2000 #sequence_revision 01-Dec-2000 #text_change 09-Jul-2004
C:Accession: F84023
R:Takami, H.; Nakasone, K.; Takaki, Y.; Maeno, G.; Sasaki, R.; Masui, N.; Fuji, F.; Hira
Nucleic Acids Res. 28, 4317-4331, 2000
A:Title: Complete genome sequence of the alkaliphilic bacterium Bacillus halodurans and
A:Reference number: A83650; MUID:20512582; PMID:11058132
A:Accession: F84023
A:Status: preliminary
A:Molecule type: DNA
A:Residues: 1-216 <STO>
A:Cross-references: UNIPROT:Q9K8L6; GB:BA000004; NID:gi0175500; PIDN:BAB067
A:Experimental source: strain C-125
C:Genetics:
A:Gene: BH2990
C:Superfamily: phoU protein

Query Match      55.1%; Score 43; DB 2; Length 216;
Best Local Similarity 57.1%; Pred. No. 5.2;
Matches 8; Conservative 3; Mismatches 3; Indels 0; Gaps 0;

Qy 1 SRLDYLKSSLLHLG 14
   | : || || | : | : |
Db 8 SLDYLKQGILMMG 21

RESULT 4
S46812
hypothetical protein YHR073w - yeast (Saccharomyces cerevisiae)
C:Species: Saccharomyces cerevisiae
C:Date: 28-Oct-1994 #sequence_revision 28-Oct-1994 #text_change 09-Jul-2004
C:Accession: S46812
R:Favellio, T.
submitted to the EMBL Data Library, June 1994
A:Description: The sequence of S. cerevisiae cosmid 9205.
A:Reference number: S46795
A:Accession: S46812
A:Molecule type: DNA
A:Residues: 1-996 <FAV>
A:Cross-references: UNIPROT:P38713; EMBL:U10556; NID:g500825; PID:g500833; GSPDB:GN000008
C:Genetics:
A:Gene: SGD:YHR073W
A:Cross-references: SGD:S0001115
A:Map position: 8R

Query Match      55.1%; Score 43; DB 2; Length 996;
Best Local Similarity 69.2%; Pred. No. 28;
Matches 9; Conservative 0; Mismatches 4; Indels 0; Gaps 0;

Qy 3 LDYLKSSLLHLGS 15
   ||| ||||| |||
Db 64 LDYTTKSLHLKGS 76

```

```

RESULT 5
C87315
hypothetical protein CC0532 [imported] - Caulobacter crescentus
C:Species: Caulobacter crescentus
C:Date: 20-Apr-2001 #sequence_revision 20-Apr-2001 #text_change 09-Jul-2004
C:Accession: C87315
R:Nierman, W.C.; Feidblyum, T.V.; Paulsen, I.T.; Nelson, K.E.; Eisen, J.; Heidelberg, J.L.
B.; Laub, M.C.; DeBoy, R.T.; Dodson, R.J.; Durkin, A.S.; Gwinn, M.L.; Haft, D.H.; Kolon
n, J.; Ermolaeva, M.; White, O.; Salzberg, S.L.; Shapiro, L.; Venter, J.C.; Fraser, C.M.
Proc. Natl. Acad. Sci. U.S.A. 98, 4136-4141, 2001
A:Title: Complete Genome Sequence of Caulobacter crescentus.
A:Reference number: A87249; MUID:21173698; PMID:11259647
A:Accession: C87315
A:Status: preliminary
A:Molecule type: DNA
A:Residues: 1-403 <STO>
A:Cross-references: UNIPROT:Q9AAR4; GB:AE005673; NID:gl3421717; PIDN:AAK22519.1; GSPDB:GN
C:Genetics:
A:Gene: CC0532

Query Match      52.6%; Score 41; DB 2; Length 403;
Best Local Similarity 50.0%; Pred. No. 23;
Matches 6; Conservative 5; Mismatches 1; Indels 0; Gaps 0;

Qy 3 LDYLKSSLLHLG 14
   ||||| : || : ||
Db 280 IDYLRAIVLHAG 291

RESULT 6
B64433
probable O-antigen transporter homolog - Methanococcus jannaschii
C:Species: Methanococcus jannaschii
C:Date: 13-Sep-1996 #sequence_revision 13-Sep-1996 #text_change 09-Jul-2004
C:Accession: B64433
R:Bult, C.J.; White, O.; Olsen, G.J.; Zhou, L.; Fleischmann, R.D.; Sutton, G.G.; Blake, A.;
Reich, C.I.; Overbeek, R.; Kirkness, E.F.; Weinatock, K.G.; Merrick, J.M.; Glodek, A.;
rson, J.D.; Sadow, P.W.; Hanna, M.C.; Cotton, M.D.; Roberts, K.M.; Hurst, M.A.
Science 273, 1058-1073, 1996
A:Authors: Kaine, B.P.; Borodovsky, M.; Klenk, H.P.; Fraser, C.M.; Smith, H.O.; Woese, C.
A:Title: Complete genome sequence of the methanogenic archaeon, Methanococcus jannaschii
A:Reference number: A64300; MUID:96337999; PMID:8688087
A:Accession: B64433
A:Status: preliminary
A:Molecule type: DNA
A:Residues: 1-507 <BUL>
A:Cross-references: UNIPROT:Q58467; GB:U67549; GB:U77117; NID:g2826363; PIDN:AAB99070.1;
C:Genetics:
A:Map position: REV1009618-1008095
A:Start codon: TTG

Query Match      52.6%; Score 41; DB 2; Length 507;
Best Local Similarity 57.1%; Pred. No. 30;
Matches 8; Conservative 4; Mismatches 2; Indels 0; Gaps 0;

Qy 1 SRLDYLKSSLLHLG 14
   ||||| ||||| : |
Db 77 NRLDMLKSSIVFVG 90

RESULT 7
D72222
conserved hypothetical protein - Thermotoga maritima (strain MSB8)
C:Species: Thermotoga maritima
C:Date: 11-Jun-1999 #sequence_revision 11-Jun-1999 #text_change 09-Jul-2004
C:Accession: D72222
R:Nelson, K.E.; Clayton, R.A.; Gill, S.R.; Gwinn, M.L.; Dodson, R.J.; Haft, D.H.; Hickey
Garrett, M.M.; Stewart, A.M.; Cotton, M.D.; Pratt, M.S.; Phillips, C.A.; Richardson, D.;
C.M.
Nature 399, 323-329, 1999
A:Title: Evidence for lateral gene transfer between Archaea and Bacteria from genome seq
A:Reference number: A72200; MUID:99287316; PMID:10360571
A:Accession: D72222

```

A:Status: preliminary
A:Molecule type: DNA
A:Residues: 1-565 <ARN>
A:Cross-references: UNIPROT:Q9X210; GB:AB001809; GB:AB000512; NID:G9482257; PIDN:AAD3674
A:Experimental source: strain MSB8
C:Genetics:
A:Gene: TM1682

Query Match 52.6%; Score 41; DB 2; Length 565;
Best Local Similarity 66.7%; Pred. No. 34;
Matches 8; Conservative 3; Mismatches 1; Indels 0; Gaps 0;

QY 3 LDYLKSSLHLG 14

|.:|:|:|:|:|:|

248 LVFLRSLHLG 259

RESULT 8

A71647

glycine-tRNA ligase (EC 6.1.1.14) (GlyS) RP849 - Rickettsia prowazekii

C:Species: Rickettsia prowazekii

C:Date: 21-Nov-1998 #sequence_revision 21-Nov-1998 #text_change 09-Jul-2004

C:Accession: A71647

R:Andersson, S.G.E.; Zomorodipour, A.; Andersson, J.O.; Sicheritz-Ponten, T.; Alenmark, U.

Nature 396, 133-140, 1998

A:Title: The genome sequence of Rickettsia prowazekii and the origin of mitochondria.

A:Reference number: A71630; MUID:99039499; PMID:9823893

A:Accession: A71647

A:Status: preliminary; nucleic acid sequence not shown; translation not shown

A:Molecule type: DNA

A:Residues: 1-648 <AND>

A:Cross-references: UNIPROT:Q9ZCB1; GB:AJ235273; GB:AJ235269; NID:G3861237; PIDN:CAA1527

A:Experimental source: strain Madrid B

C:Genetics:

A:Gene: GlyS; RP849

C:Superfamily: glycine-tRNA ligase beta chain

C:Keywords: aminoacyl-tRNA synthetase; ligase; protein biosynthesis

Query Match 52.6%; Score 41; DB 2; Length 648;

Best Local Similarity 64.7%; Pred. No. 39;

Matches 11; Conservative 1; Mismatches 3; Indels 2; Gaps 1;

QY 1 SRLDYKSSLH--LGS 15

|:|:|:|:|:|

317 SRLDKLSVIFHTKLS 333

RESULT 9

H84605

hypothetical protein At2g21850 [imported] - Arabidopsis thaliana

C:Species: Arabidopsis thaliana (mouse-ear cress)

C:Date: 02-Feb-2001 #sequence_revision 02-Feb-2001 #text_change 09-Jul-2004

C:Accession: H84605

R:Lin, X.; Kaul, S.; Rounsley, S.D.; Shea, T.P.; Benito, M.I.; Town, C.D.; Fujii, C.Y.;

M.; Koo, H.; Moffat, K.S.; Cronin, L.A.; Shen, M.; VanAken, S.B.; Umayam, L.; Tallon, L.;

eusa, D.; Nierman, W.C.; White, O.; Eisen, J.A.; Salzberg, S.L.; Fraser, C.M.; Venter, J.

Nature 402, 761-768, 1999

A:Title: Sequence and analysis of chromosome 2 of the plant Arabidopsis thaliana.

A:Reference number: A84420; MUID:20083487; PMID:10617197

A:Accession: H84605

A:Status: preliminary

A:Molecule type: DNA

A:Residues: 1-772 <DNA>

A:Cross-references: UNIPROT:Q9SJ14; GB:AB002093; NID:g4417297; PIDN:AAD20422.1; GSPDGB:GN

A:Gene: At2g21850

A:Map position: 2

Query Match 52.6%; Score 41; DB 2; Length 772;

Best Local Similarity 58.3%; Pred. No. 48;

Matches 7; Conservative 3; Mismatches 2; Indels 0; Gaps 0;

QY 2 RLDYKSSLHL 13

|.:|:|:|:|:|

2 RLDYKSSLHL 13

Db 189 RLDYTRTSLCHI 200

RESULT 10

S33918

dynamitin-like protein MCM1 - yeast (Saccharomyces cerevisiae)

N:Alternate names: protein OS010; protein YOR211c; protein YOR50-1

C:Species: Saccharomyces cerevisiae

C:Date: 31-Dec-1993 #sequence_revision 31-Dec-1993 #text_change 09-Jul-2004

C:Accession: S33918; S36728; A42137; S48883; S60938; S67104; S67103; S71713

R:Guan, K.; Farh, L.; Marshhall, T.K.; Deschenes, R.J.

Curr. Genet. 24, 141-148, 1993

A:Title: Normal mitochondrial structure and genome maintenance in yeast requires the dyna

A:Reference number: S33918; MUID:93365024; PMID:7916673

A:Accession: S33918

A:Molecule type: DNA

A:Residues: 1-902 <GUA1>

A:Cross-references: UNIPROT:P32266; EMBL:L07419

A>Note: It is uncertain whether Met-1 or Met-60 is the initiator

R:Guan, K.; Farh, L.; Marshhall, T.; Deschenes, R.J.

submitted to the EMBL Data Library, January 1992

A:Reference number: S36728

A:Accession: S36728

A:Molecule type: DNA

A:Residues: 1-190, 'A', 192-902 <GUA2>

A:Cross-references: EMBL:L07419

R:Jones, B.A.; Fangman, W.L.

Genes Dev. 6, 380-389, 1992

A:Title: Mitochondrial DNA maintenance in yeast requires a protein containing a region re

A:Reference number: A42137; MUID:92192451; PMID:1532158

A:Accession: A42137

A:Molecule type: DNA

A:Residues: 60-149, 'C', 151-902 <JON>

A:Cross-references: EMBL:X62834; NID:G3956; PIDN:CAA44637.1; PID:G3957

A>Note: sequence extracted from NCBI backbone (NCBIN:88065, NCBI:P:88066)

R:Kalo, D.; Carles, C.; Sentenac, A.; Thuriaux, P.

Proc. Natl. Acad. Sci. U.S.A. 90, 5524-5528, 1993

A:Title: Interactions between three common subunits of yeast RNA polymerases I and III.

A:Reference number: S48883; MUID:93296170; PMID:8516295

A:Accession: S48883

A:Status: nucleic acid sequence not shown; translation not shown

A:Molecule type: DNA

A:Residues: 734-902 <LAL>

A:Cross-references: EMBL:L11274; NID:G295632; PIDN:AAB59316.1; PID:G295633

A>Note: the nucleotide sequence was submitted to the EMBL Data Library, February 1993

R:Gallison, F.; Dujon, B.

submitted to the EMBL Data Library, October 1995

A:Description: Sequence and analysis of a 33 kb fragment from the right arm of chromosome

A:Reference number: S60938

A:Accession: S60938

A:Molecule type: DNA

A:Residues: 1-805 <GAL>

A:Cross-references: EMBL:X92441; NID:G1050762; PIDN:CAA63174.1; PID:G1050763

R:Boyer, J.; Fairhead, C.; Gallon, L.; Gallison, F.; Michaux, G.; Thierry, A.; Dujon, B.

submitted to the Protein Sequence Database, July 1996

A:Reference number: S67104

A:Accession: S67104

A:Molecule type: DNA

A:Residues: 1-805 <BOY>

A:Cross-references: EMBL:275119; MIPS:YOR211c

A:Experimental source: strain S288C

R:Hughes, B.; Pohl, T.M.

submitted to the Protein Sequence Database, July 1996

A:Reference number: S66685

A:Accession: S67103

A:Molecule type: DNA

A:Residues: 179-902 <HUG>

A:Cross-references: EMBL:275119; MIPS:YOR211c

A:Experimental source: strain S288C

R:Gallison, F.; Dujon, B.

Yeast 12, 877-885, 1996

A:Title: Sequence and analysis of a 33 kb fragment from the right arm of chromosome XV of

A;Reference number: S71713; MUID:96437977; PMID:8840505
 A;Accession: S71713
 A;Molecule type: DNA
 A;Residues: 1-150 <GAW>
 A;Cross-references: EMBL:X92441
 C;Genetics:
 A;Gene: SGD:MGM1
 A;Cross-references: SGD:S0005737; MIPS:YOR211C
 A;Map position: 15R
 A;Genome: nuclear
 C;Function:

A;Description: mitochondrial genome maintenance
 C;Keywords: GTP binding; mitochondrion; nucleotide binding; P-loop
 F;238-245/Region: nucleotide-binding motif A (P-loop)

Query Match 52.6%; Score 41; DB 2; Length 902;
 Best Local Similarity 66.7%; Pred. No. 56;
 Matches 8; Conservative 2; Mismatches 2; Indels 0; Gaps 0;

QY 5 YLKSSLLHLGSR 16
 |::|::|::|::|
 Db 48 YIKSPLVHLHSR 59

RESULT 11

T17297
 hypothetical protein DKFZp586F1023.1 - human
 C;Species: Homo sapiens (man)
 C;Date: 15-Oct-1999 #sequence_revision 15-Oct-1999 #text_change 16-Aug-2004
 C;Accession: T17297
 R;Ansoorge, W.; Wirkner, U.; Mewes, H.W.; Gassenhuber, J.; Wiemann, S.
 submitted to the Protein Sequence Database, September 1999
 A;Reference number: Z18729

A;Accession: T17297
 A;Status: preliminary
 A;Molecule type: mRNA
 A;Residues: 1-960 <ANS>
 A;Cross-references: UNIPROT:Q9Y4C8; EMBL:AL117547
 A;Experimental source: adult uterus; clone DKFZp586F1023
 C;Genetics:

A;Note: DKFZp586F1023.1
 C;Superfamily: ribonucleoprotein repeat homology

Query Match 52.6%; Score 41; DB 2; Length 960;
 Best Local Similarity 53.3%; Pred. No. 60;
 Matches 8; Conservative 3; Mismatches 4; Indels 0; Gaps 0;

QY 1 SRLDYLKSSLLHLGS 15
 |::|::|::|::|
 Db 206 SDMDYLKSKMVKAGS 220

RESULT 12

AF3460
 histidinol dehydrogenase (EC 1.1.1.23) [imported] - Brucella melitensis (strain 16M)
 C;Species: Brucella melitensis
 C;Date: 01-Feb-2002 #sequence_revision 01-Feb-2002 #text_change 16-Aug-2004
 C;Accession: AF3460
 R;DeVecchio, V.G.; Kapatal, V.; Redkar, R.J.; Patra, G.; Mujar, C.; Los, T.; Ivanova,
 ; Mazur, M.; Goltzman, E.; Selkov, E.; Elzer, P.H.; Hagius, S.; O'Callaghan, D.; Letess
 proc. Natl. Acad. Sci. U.S.A. 99, 443-448, 2002
 A;Title: The genome sequence of the facultative intracellular pathogen Brucella melitensis
 A;Reference number: AD3252; PMID:11756688

A;Accession: AF3460
 A;Status: preliminary
 A;Molecule type: DNA
 A;Residues: 1-430 <KUR>
 A;Cross-references: UNIPROT:Q8YF59; GB:AE008917; PIDN:AAU52849.1; PID:g17983691; GSPDB:G
 A;Experimental source: strain 16M
 C;Genetics:

A;Gene: BME1668
 A;Map position: 1
 C;Superfamily: Histidinol dehydrogenase; histidinol dehydrogenase homology

C;Keywords: oxidoreductase

Query Match 51.9%; Score 40.5; DB 2; Length 430;
 Best Local Similarity 68.8%; Pred. No. 31;
 Matches 11; Conservative 2; Mismatches 2; Indels 1; Gaps 1;

QY 1 SRLDYLK-SSLHLGS 15
 |::|::|::|::|
 Db 383 SVLDYMKRTSLKLG 398

RESULT 13

DB3980
 hypothetical protein BH2644 [imported] - Bacillus halodurans (strain C-125)
 C;Species: Bacillus halodurans
 C;Date: 01-Dec-2000 #sequence_revision 01-Dec-2000 #text_change 09-Jul-2004
 C;Accession: DB3980
 R;Takami, H.; Nakasone, K.; Takaki, Y.; Maeno, G.; Sasaki, R.; Masui, N.; Fujii, F.; Hiran
 Nucleic Acids Res. 28, 4317-4331, 2000
 A;Title: Complete genome sequence of the alkaliphilic bacterium Bacillus halodurans and
 A;Reference number: A83650; MUID:20512582; PMID:11058132

A;Accession: DB3980
 A;Status: preliminary
 A;Molecule type: DNA
 A;Residues: 1-354 <STO>
 A;Cross-references: UNIPROT:Q9K9K2; GB:AP001516; GB:BA000004; NID:g10175192; PIDN:BA063
 A;Experimental source: strain C-125
 C;Genetics:
 A;Gene: BH2644

Query Match 51.3%; Score 40; DB 2; Length 354;
 Best Local Similarity 53.8%; Pred. No. 31;
 Matches 7; Conservative 3; Mismatches 3; Indels 0; Gaps 0;

QY 2 RLDYKSSLLHLG 14
 |::|::|::|::|
 Db 233 KLDYVYKALAHLG 245

RESULT 14

DB6442
 probable amino acid permease [imported] - Arabidopsis thaliana
 C;Species: Arabidopsis thaliana (mouse-ear cress)
 C;Date: 02-Mar-2001 #sequence_revision 02-Mar-2001 #text_change 09-Jul-2004
 C;Accession: DB6442
 R;Theologis, A.; Ecker, J.R.; Palm, C.J.; Federspiel, N.A.; Kaul, S.; White, O.; Alonso,
 Chin, C.W.; Chung, M.K.; Conn, L.; Conway, A.B.; Conway, A.R.; Creasy, T.H.; Dewar, K.;
 ansen, N.F.; Hughes, B.; Huizar, L.
 Nature 408, 816-820, 2000

A;Authors: Hunter, J.L.; Jenkins, J.; Johnson-Hopson, C.; Khan, S.; Khaykin, E.; Kim, C.-
 C.A.; Li, J.H.; Li, Y.; Lin, X.; Liu, S.X.; Liu, Z.A.; Luros, J.S.; Maiti, R.; Marziali,
 Rizzo, M.; Rooney, T.; Rowley, D.; Sakano, H.
 A;Authors: Salzberg, S.L.; Schwartz, J.R.; Shinn, P.; Southwick, A.M.; Sun, H.; Tallon,
 ker, M.; Wu, D.; Yu, G.; Fraser, C.M.; Venter, J.C.; Davis, R.W.
 A;Title: Sequence and analysis of chromosome 1 of the plant Arabidopsis.
 A;Reference number: A86141; MUID:21016719; PMID:11130712

A;Accession: DB6442
 A;Status: preliminary
 A;Molecule type: DNA
 A;Residues: 1-495 <STO>
 A;Cross-references: UNIPROT:Q9C6S5; GB:AE005172; NID:g11136719; PIDN:AAG31300.1; GSPDB:G
 C;Genetics:
 A;Map position: 1

Query Match 51.3%; Score 40; DB 2; Length 495;
 Best Local Similarity 69.2%; Pred. No. 44;
 Matches 9; Conservative 1; Mismatches 3; Indels 0; Gaps 0;

QY 3 LDYKSSLLHLGS 15
 |::|::|::|::|
 Db 149 LDYKSGVPALGS 161

RESULT 15

T37738
 hypothetical protein SPAC167.05 - fission yeast (Schizosaccharomyces pombe)
 C:Species: Schizosaccharomyces pombe
 C:Date: 03-Dec-1999 #sequence_revision 03-Dec-1999 #text_change 09-Jul-2004
 C:Accession: T37738; T38953
 R:Rieger, M.; Wood, V.; Rajandream, M.A.; Barrell, B.G.
 submitted to the EMBL Data Library, January 1999
 A:Reference number: Z21743
 A:Accession: T37738
 A:Molecule type: DNA
 A:Residues: 1-601 <RIE>
 A:Cross-references: UNIPROT:P87132; EMBL:AL035248; PIDN:CAA22850.1; GSPDB:GN000066; SPDB:
 R:Skelton, J.; Churcher, C.M.; Barrell, B.G.; Rajandream, M.A.; Wood, V.
 submitted to the EMBL Data Library, May 1997
 A:Reference number: Z21819
 A:Accession: T38953
 A:Molecule type: DNA
 A:Residues: 471-601 <SKE>
 A:Cross-references: EMBL:Z95396; PIDN:CAB08759.1; GSPDB:GN000066; SPDB:SPAC57A7.01
 A:Experimental source: strain 972h-; cosmid c57A7
 C:Genetics:
 A:Gene: SPDB:SPAC57A7.01; SPDB:SPAC167.05
 A:Map position: 1

Query Match 51.3%; Score 40; DB 2; Length 601;
 Best Local Similarity 50.0%; Pred. No. 55;
 Matches 7; Conservative 5; Mismatches 2; Indels 0; Gaps 0;

Qy 3 LDYKSSLLHLGSR 16

Db 527 IDYIEPSLVVMGSR 540

Search completed: November 15, 2005, 12:23:30
 Job time : 16.3158 secs

This Page Blank (uspto)

GenCore version 5.1.6
Copyright (c) 1993 - 2005 Compugen Ltd.

OM protein - protein search, using sw model

Run on: November 15, 2005, 12:12:49 ; Search time 116.632 Seconds
(without alignments)
70.249 Million cell updates/sec

Title: US-10-820-998-1

Perfect score: 78

Sequence: 1 SRLDYKSSLLHLGSR 16

Scoring table:

BLOSUM62

Gapop 10.0 , Gapext 0.5

Searched: 1612378 seqs, 512079187 residues

Total number of hits satisfying chosen parameters: 1612378

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database : Uniprot 03.*

1: uniprot_sprot.*

2: uniprot_trembl.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match %	Length	DB ID	Description
1	49	62.8	526	2	Q8TGO0
2	49	62.8	603	2	Q8IP91
3	49	62.8	603	2	Q8T471
4	48	61.5	435	2	Q19242
5	44	56.4	88	2	Q87CX6
6	44	56.4	715	2	Q7Z4M8
7	44	56.4	847	2	Q9AZU5
8	43	55.1	120	2	Q6LKE8
9	43	55.1	216	2	Q9K8L6
10	43	55.1	522	2	Q7MSA0
11	43	55.1	720	2	Q9INT4
12	43	55.1	996	1	Q8H3 YEAST
13	42	53.8	140	2	Q8OUJ4
14	42	53.8	350	2	Q8IYV9
15	42	53.8	436	2	Q9NLE8
16	42	53.8	491	2	Q8N115
17	42	53.8	505	2	Q8NA69
18	42	53.8	558	2	Q8N4S9
19	42	53.8	726	2	Q6LST8
20	41	52.6	82	2	Q63E26
21	41	52.6	139	2	Q6FPP8
22	41	52.6	204	2	Q6RG05
23	41	52.6	403	2	Q9AAR4
24	41	52.6	507	1	YAE8_METJA
25	41	52.6	565	2	Q9X210
26	41	52.6	653	2	Q6F9W0
27	41	52.6	665	1	SYGB RICPR
28	41	52.6	754	2	P79816
29	41	52.6	772	2	Q9S114
30	41	52.6	791	2	Q9L3F5
31	41	52.6	791	2	Q9K496

32 41 52.6 793 2 Q9EWB2 streptomyc
33 41 52.6 793 2 Q82IV7 streptomyc
34 41 52.6 804 2 Q9EWB5 streptomyc
35 41 52.6 805 2 Q9L383 streptomyc
36 41 52.6 823 2 Q9VFR5 drosophila
37 41 52.6 876 2 Q6NNY2 drosophila
38 41 52.6 881 1 MGML YEAST
39 41 52.6 915 2 Q7XAV6 saccharomyc
40 41 52.6 960 1 K682 HUMAN
41 41 52.6 960 2 Q9BPY6 homo sapien
42 41 52.6 4133 2 Q63NJ3 burkholderi
43 41 52.6 7523 2 Q6D738 erwinia car
44 40.5 51.9 430 1 HISX_BRUME bruceella me
45 40.5 51.9 430 1 HISX_BRUSO bruceella su

ALIGNMENTS

RESULT 1

Q8TGO0 PRELIMINARY; PRT; 526 AA.
AC Q8TGO0;
DT 01-JUN-2002 (TrEMBLrel. 21, Created)
DT 01-JUN-2002 (TrEMBLrel. 21, Last sequence update)
DT 01-MAR-2004 (TrEMBLrel. 26, Last annotation update)
DE LD05405P.
GN ORFNames=CG31759;
OS Drosophila melanogaster (fruit fly).
OC Eukaryota; Metazoa; Arthropoda; Hexapoda; Insecta; Pterygota;
OC Neoptera; Endopterygota; Diptera; Brachycera; Muscomorpha;
OC Ephydroidea; Drosophilidae; Drosophila.
OX NCBI_TaxID=7227;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=Berkley;
RA Stapleton M., Brokstein P., Hong L., Agbayani A., Carlson J.,
RA Champe M., Chavez C., Dorsett V., Farfan D., Frise E., George R.,
RA Gonzalez M., Guarin H., Li P., Liao G., Miranda A., Mungall C.J.,
RA Nuoo J., Pacleb J., Paragas V., Park S., Phouanavong S., Wan K.,
RA Yu C., Lewis S.E., Rubin G.M., Celnik S.
RL Submitted (DEC-2001) to the EMBL/GenBank/DBJ databases.
DR EMBL; AF069343; AAL39488.1; -.
DR FlyBase; FBgn0051759; CG31759.
DR InterPro; IPR005135; Exo endo phos.
DR InterPro; IPR002052; N6_Mtase.
DR Pfam; PF033372; Exo endo phos; 1.
DR PROSITE; PS00092; N6_MTASE; UNKNOWN 1.
SQ SEQUENCE 526 AA; 59525 MW; 6FB4E6D18737F798 CRC64;

Query Match 62.8%; Score 49; DB 2; Length 526;

Best Local Similarity 66.7%; Pred. NO. 6.5;

Matches 10; Conservative 1; Mismatches 4; Indels 0; Gaps 0;

QY 1 SRLDYKSSLLHLGHS 15

DB 297 SRFDLLDSQLHLGS 311

RESULT 2

Q8IP91 PRELIMINARY; PRT; 603 AA.
AC Q8IP91;
DT 01-MAR-2003 (TrEMBLrel. 23, Created)
DT 01-MAR-2003 (TrEMBLrel. 23, Last sequence update)
DT 01-MAR-2004 (TrEMBLrel. 26, Last annotation update)
DE CG31759-PA (CG31759-pb).
GN ORFNames=CG31759;
OS Drosophila melanogaster (fruit fly).
OC Eukaryota; Metazoa; Arthropoda; Hexapoda; Insecta; Pterygota;
OC Neoptera; Endopterygota; Diptera; Brachycera; Muscomorpha;
OC Ephydroidea; Drosophilidae; Drosophila.
OX NCBI_TaxID=7227;

RN SEQUENCE FROM N.A.
 RP MEDLINE=20196006; PubMed=10731132; DOI=10.1126/science.287.5461.2185;
 RA Adams M.D., Celniker S.E., Holt R.A., Evans C.A., Gocayne J.D.,
 RA Anantadesi P.G., Scherer S.E., Li P.W., Hoskins R.A., Galle R.F.,
 RA George R.A., Lewis S.E., Richards S., Ashburner M., Henderson S.N.,
 RA Sutton G.G., Wortman J.R., Yandell M.D., Zhang Q., Chen L.X.,
 RA Brandon R.C., Rogers Y.H., Blaziej R.G., Champe M., Pfeiffer B.D.,
 RA Wan K.H., Doyle C., Baxter E.G., Helt G., Nelson C.R., Gabor G.L.,
 RA April J.F., Agbayani A., An H.J., Andrews-Pfankoch C., Baldwin D.,
 RA Ballew R.M., Basu A., Baxendale J., Bayraktaroglu L., Beasley E.M.,
 RA Beeson K.Y., Benos P.V., Berman B.P., Bhandari D., Bolshakov S.,
 RA Borkova D., Botchan M.R., Bouck J., Brokstein P., Brottier P.,
 RA Burtis K.C., Busam D.A., Butler H., Cadieu E., Center A., Chandra I.,
 RA Cherry J.M., Cawley S., Dahle C., Davenport L.B., Davies P.,
 RA de Pablo B., Delcher A., Deng Z., Mays A.D., Dew I., Dietz S.M.,
 RA Dodson K., Doup L.E., Downes M., Dugan-Rocha S., Dunkov B.C., Dunn P.,
 RA Durbin K.J., Evangelista C.C., Ferraz C., Ferreira S., Fleischmann W.,
 RA Foster C., Gabrielian A.E., Garg N.S., Gelbart W.M., Glasser K.,
 RA Glodok A., Gong F., Gorrell J.H., Gu Z., Guan P., Harris M.,
 RA Harris N.L., Harvey D., Heiman T.J., Hernandez J.R., Houck J.,
 RA Hestlin D., Houston K.A., Howland T.J., Wei M.H., Ibegwan C.,
 RA Jalali M., Kalush F., Karpen G.H., Ke Z., Kennison J.A., Ketchum K.A.,
 RA Kimmel B.E., Kodira C.D., Kraft C., Kravitz S., Kulp D., Lai Z.,
 RA Lamko P., Lei Y., Levitsky A.A., Li J., Li Z., Liang Y., Lin X.,
 RA Liu X., Mattei B., McIntosh T.C., McLeod M.P., McPherson D.,
 RA Markulov G., Milshina N.V., Mobarry C., Morris J., Moshrefi A.,
 RA Mount S.M., Moy M., Murphy B., Murphy L., Muzny D.M., Nelson D.L.,
 RA Nelson D.R., Nelson K.A., Nixon K., Nuskern D.R., Pacleb J.M.,
 RA Palazzolo M., Pittman G.S., Pan S., Pollard J., Puri V., Reese M.G.,
 RA Reinert K., Remington K., Saunders R.D., Scheeler F., Shen H.,
 RA Shue B.C., Siden-Klamis I., Simpson M., Skupski M.P., Smith T.,
 RA Spier E., Spradling A.C., Stapleton M., Strong R., Sun B.,
 RA Svirskas R., Tector C., Turner R., Venter E., Wang A.H., Wang X.,
 RA Wang Z.Y., Wassarman D.A., Weinstock G.M., Weissbach J.,
 RA Williams S.M., Woodage, Worley K.C., Wu D., Yang S., Yao Q.A., Ye J.,
 RA Yeh R.F., Zaveri J.S., Zhan M., Zhang G., Zhao Q., Zheng L.,
 RA Zheng X.H., Zhong F.N., Zhong W., Zhou X., Zhu S., Zhu X., Smith H.O.,
 RA Gibbs R.A., Myers E.W., Rubin G.M., Venter J.C.;
 RT "The genome sequence of *Drosophila melanogaster*.";
 RN Science 287:2185-2195(2000).
 RN [2]
 RN SEQUENCE FROM N.A.
 RP MEDLINE=2242605; PubMed=12537568;
 RA Celniker S.E., Wheeler D.A., Kronmiller B., Carlson J.W., Halpern A.,
 RA Patel S., Adams M., Champe M., Dugan S.P., Frise E., Hodgson A.,
 RA George R.A., Hoskins R.A., Lavery T., Muzny D.M., Nelson C.R.,
 RA Pacleb J.M., Park S., Pfeiffer B.D., Richards S., Sodergren E.J.,
 RA Svirskas R., Tabor P.E., Wan K., Stapleton M., Sutton G.G., Venter C.,
 RA Weinstock G., Scher S.E., Myers E.W., Gibbs R.A., Rubin G.M.;
 RT "Finishing a whole-genome shotgun: Release 3 of the *Drosophila*
 RT melanogaster euchromatic genome sequence.";
 RN Genome Biol. 3:RESEARCH0079-RESEARCH0079(2002).
 RN [3]
 RN SEQUENCE FROM N.A.
 RP MEDLINE=22426070; PubMed=12537573;
 RA Kaminker J.S., Bergman C.M., Kronmiller B., Carlson J., Svirskas R.,
 RA Patel S., Frise E., Wheeler D.A., Lewis S.E., Rubin G.M.,
 RA Ashburner M., Celniker S.E.;
 RT "The transposable elements of the *Drosophila melanogaster* euchromatin:
 RT a genomics perspective.";
 RN Genome Biol. 3:RESEARCH0084-RESEARCH0084(2002).
 RN [4]
 RN SEQUENCE FROM N.A.
 RP MEDLINE=22426069; PubMed=12537572;
 RA Misra S., Crosby M.A., Mungall C.J., Matthews B.B., Campbell K.S.,
 RA Hradecky P., Huang Y., Kaminker J.S., Millburn G.H., Prochnik S.E.,
 RA Smith C.D., Tupy J.L., Whitfield E.J., Bayraktaroglu L., Berman B.P.,
 RA Bertencourt B.R., Celniker S.E., de Grey A.D., Drysdale R.A.,
 RA Harris N.L., Richter J., Ruseo S., Schroeder A.J., Shu S.Q.,
 RA Stapleton M., Yamada C., Ashburner M., Gelbart W.M., Rubin G.M.,
 RA Lewis S.E.;
 RT "Annotation of the *Drosophila melanogaster* euchromatic genome: a

RT systematic review.";
 RL Genome Biol. 3:RESEARCH0083-RESEARCH0083(2002).
 RN [5]
 RN SEQUENCE FROM N.A.
 RP FlyBase;
 RG Submitted (SEP-2002) to the EMBL/GenBank/DBJ databases.
 RN [6]
 RN SEQUENCE FROM N.A.
 RG FlyBase;
 RL Submitted (MAR-2004) to the EMBL/GenBank/DBJ databases.
 DR ENBL; AS003635; AAN10808.1; -;
 DR FlyBase; FBgn0051759; CG31759.
 DR InterPro; IPR005135; Exo_endo_phos.
 DR InterPro; IPR002052; N6_Mtase.
 DR Pfam; PF03372; Exo_endo_phos; 1.
 DR PROSITE; PS00092; N6_MTASE; UNKNOWN 1.
 SQ SEQUENCE 603 AA; 58498 MW; 95E241EADF204D01 CRC64;
 Query Match 62.8%; Score 49; DB 2; Length 603;
 Best Local Similarity 66.7%; Pred. No. 7.6;
 Matches 10; Conservative 1; Mismatches 4; Indels 0; Gaps 0;
 QY 1 SRLDYLKSSLLHLGS 15
 DB 374 SRFDLLDSQILHLGS 388
 RESULT 3
 Q8T471 PRELIMINARY; PRT; 603 AA.
 ID Q8T471
 AC Q8T471;
 DT 01-JUN-2002 (TrEMBLrel. 21, Created)
 DT 01-JUN-2002 (TrEMBLrel. 21, Last sequence update)
 DT 01-MAR-2004 (TrEMBLrel. 26, Last annotation update)
 DE AT13596P.
 GN ORFNames=CG31759;
 OS *Drosophila melanogaster* (fruit fly).
 OC Eukaryota; Metazoa; Arthropoda; Hexapoda; Insecta; Pterygota;
 OC Neoptera; Endopterygota; Diptera; Brachycera; Muscomorpha;
 OC Ephydroidea; Drosophilidae; Drosophila.
 OX NCBI_TaxID=7227;
 RN [1]
 RN SEQUENCE FROM N.A.
 RA Stapleton M., Brokstein P., Hong L., Agbayani A., Carlson J.,
 RA Champe M., Chavez C., Dorsett V., Dresnek D., Farfan D., Frise E.,
 RA George R., Gonzalez M., Guarin H., Kronmiller B., Li P., Liao G.,
 RA Miranda A., Mungall C.J., Nunoo J., Pacleb J., Paragas V., Park S.,
 RA Patel S., Phouanavong S., Wan K., Yu C., Lewis S.E., Rubin G.M.,
 RA Celniker S.;
 RL Submitted (MAR-2002) to the EMBL/GenBank/DBJ databases.
 DR ENBL; AY089324; AAL90062.1; -;
 DR FlyBase; FBgn0051759; CG31759.
 DR InterPro; IPR005135; Exo_endo_phos.
 DR InterPro; IPR002052; N6_Mtase.
 DR Pfam; PF03372; Exo_endo_phos; 1.
 DR PROSITE; PS00092; N6_MTASE; UNKNOWN 1.
 SQ SEQUENCE 603 AA; 58507 MW; 4FS490AB6AADBA5C CRC64;
 Query Match 62.8%; Score 49; DB 2; Length 603;
 Best Local Similarity 66.7%; Pred. No. 7.6;
 Matches 10; Conservative 1; Mismatches 4; Indels 0; Gaps 0;
 QY 1 SRLDYLKSSLLHLGS 15
 DB 374 SRFDLLDSQILHLGS 388
 RESULT 4
 Q19242 PRELIMINARY; PRT; 435 AA.
 ID Q19242
 AC Q19242; Q22274;
 DT 01-NOV-1996 (TrEMBLrel. 01, Created)
 DT 01-JUN-2003 (TrEMBLrel. 24, Last sequence update)

```

DT 25-OCT-2004 (TrEMBLrel. 28, Last annotation update)
DE Hypothetical protein T06H11.4.
GN Name=moc-1; ORFNames=T06H11.4;
OS Caenorhabditis elegans.
OC Eukaryota; Metazoa; Nematoda; Chromadorea; Rhabditida; Rhabditoidea;
OC Rhabditidae; Peloderinae; Caenorhabditis.
OX NCBI_TaxID=6239;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=Bristol N2;
RA none;
RX MEDLINE=99069613; PubMed=9851916;
RT "Genome sequence of the nematode C.elegans: A platform for
RL investigating biology.";
RL Science 282:2012-2018 (1998).
RN [2]
RP SEQUENCE FROM N.A.
RC STRAIN=Bristol N2;
RA Kershaw J.K.;
RL Submitted (JUN-1995) to the EMBL/GenBank/DBJ databases.
DR EMBL; Z49887; CAA90060.2; -.
DR EMBL; Z49889; CAA90060.2; JOINED.
DR EMBL; Z49887; CAA90069.2; JOINED.
DR EMBL; Z49889; CAA90069.2; -.
DR FIK; T20638; T20638.
DR HSSP; P12281; IFC5.
DR WormBase; WBGene0003384; moc-1.
DR WormPeP; T06H11.4; CE31596.
DR GO; GO:0006777; P:Mo-nolymphoprotein cofactor biosynthesis; IEA.
DR InterPro; IPR001453; MoCF_biosynth.
DR InterPro; IPR008285; MoCF_bios_C.
DR InterPro; IPR005111; Moea_C.
DR InterPro; IPR005110; Moea_N.
DR Pfam; PF00994; MoCF_biosynth; 1.
DR Pfam; PF03454; Moea_C; 1.
DR Pfam; PF03453; Moea_N; 1.
DR ProDom; PD002460; MoCF_biosynth; 1.
DR TIGRFAMs; TIGR00177; molyb_syn; 1.
DR PROSITE; PS01079; MoCF_BIOSYNTHESIS_2; UNKNOWN_1.
KW Hypothetical protein.
SQ SEQUENCE 435 AA; 47077 MW; 5DAE9CDB59615D99 CRC64;

Query Match 61.5%; Score 48; DB 2; Length 435;
Best Local Similarity 72.7%; Pred. No. 8;
Matches 8; Conservative 3; Mismatches 0; Indels 0; Gaps 0;

QY 4 DYKSSLLHLG 14
DB 273 DYMKALLHLG 283

RESULT 5
Q87CX6 PRELIMINARY; PRT; 88 AA.
AC Q87CX6;
DT 01-JUN-2003 (TrEMBLrel. 24, Created)
DT 01-JUN-2003 (TrEMBLrel. 24, Last sequence update)
DT 01-JUN-2003 (TrEMBLrel. 24, Last annotation update)
DE Hypothetical protein.
GN OrderedLocustNames=PD0921;
OS Xylella fastidiosa (strain Temecula / ATCC 700964).
OC Bacteria; Proteobacteria; Gammaproteobacteria; Xanthomonadales;
OC Xanthomonadaceae; Xylella.
OX NCBI_TaxID=183190;
RN [1]
RP SEQUENCE FROM N.A.
RX MEDLINE=22421331; PubMed=12533478;
RX DOI=10.1128/JB.185.3.1018-1026.2003;
RA Van Sluys M.A., de Oliveira M.C., Monteiro-Vitorello C.B., Moon D.H.,
RA Miyaki C.Y., Furlan L.R., Camargo L.E.A., da Silva A.C.R., Moon D.H.,
RA Takita M.A., Lemos E.G.M., Machado M.A., Ferro M.I.T., da Silva F.R.,
RA Goldman M.H.S., Goldman G.H., Lemos M.V.F., El-Dorry H., Tsai S.M.,
RA Carrer H., Carraro D.M., de Oliveira R.C., Nunes L.R., Siqueira W.J.,

RA Coutinho L.L., Kimura E.T., Ferro E.S., Harakava R., Kuramae E.E.,
RA Marino C.L., Gigliotti E., Abreu I.B., Alves L.M.C., do Amaral A.M.,
RA Baia G.S., Blanco S.R., Brito M.S., Cannavan F.S., Celestino A.V.,
RA da Cunha A.F., Fenille R.C., Ferro J.A., Formighieri E.F., Kishi L.T.,
RA Leoni S.G., Oliveira A.R., Rosa V.B. Jr., Sasaki F.T., Sena J.A.D.,
RA de Souza A.A., Truffi D., Takumo F., Yanai G.M., Zaros L.G.,
RA Civerolo E.L., Simpson A.J.G., Almeida N.F. Jr., Setubal J.C.,
RA Kitajima J.F.;
RT "Comparative analyses of the complete genome sequences of Pierce's
RT disease and citrus variegated chlorosis strains of Xylella
RT fastidiosa.";
RL J. Bacteriol. 185:1018-1026(2003).
DR EMBL; AB012556; AAO28786.1; -.
KW Complete proteome; Hypothetical protein.
SQ SEQUENCE 88 AA; 9740 MW; EB7B6B6BC0D540F3 CRC64;

Query Match 56.4%; Score 44; DB 2; Length 88;
Best Local Similarity 66.7%; Pred. No. 7.3;
Matches 10; Conservative 0; Mismatches 5; Indels 0; Gaps 0;

QY 1 SLDYKSSLLHLGS 15
DB 70 SELCYLKRLHLSS 84

RESULT 6
Q724M8 PRELIMINARY; PRT; 715 AA.
AC Q724M8;
DT 01-OCT-2003 (TrEMBLrel. 25, Created)
DT 01-OCT-2003 (TrEMBLrel. 25, Last sequence update)
DT 01-MAR-2004 (TrEMBLrel. 26, Last annotation update)
DE MTRR protein (Fragment).
GN Name=MTRR;
OS Homo sapiens (Human).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
OX NCBI_TaxID=9606;
RN [1]
RP SEQUENCE FROM N.A.
RC TISSUE=Lung;
RX MEDLINE=22388257; PubMed=12477932; DOI=10.1073/pnas.242603899;
RA Strausberg R.L., Feingold E.A., Grouse L.H., Derge J.G.,
RA Klausner R.D., Collins F.S., Wagner L., Shenmen C.M., Schuler G.D.,
RA Altschul S.F., Zeeberg B., Buetow K.H., Schaefer C.F., Bhat N.K.,
RA Hopkins R.F., Jordan H., Moore T., Max S.I., Wang J., Hsieh F.,
RA Diatchenko L., Marusina K., Farmer A.A., Rubin G.M., Hong L.,
RA Stapleton M., Soares M.B., Bonaldo M.F., Casavant T.L., Scheetz T.E.,
RA Brownstein M.J., Usdin T.B., Toshiyuki S., Carninci P., Prange C.,
RA Raha S., Loquellano N.A., Peters G.J., Abramson R.D., Mullaly S.J.,
RA Bosak S.A., McEwan P.J., McKernan K.J., Malek J.A., Gunaratne P.H.,
RA Richards S., Worley K.C., Hale S., Garcia A.M., Gay L.J., Hulyk S.W.,
RA Villalon D.K., Muzny D.M., Sodergren E.J., Lu X., Gibbs R.A.,
RA Fahey J., Helton E., Kettelman M., Maman A., Rodriguez S., Sanchez A.,
RA Whiting M., Maman A., Young A.C., Shevchenko Y., Bouffard G.G.,
RA Blakesley R.W., Touchman J.W., Green E.D., Dickson M.C.,
RA Rodriguez A.C., Grimwood J., Schmutz J., Myers R.M., Butterfield Y.S.,
RA Krzywinski M.I., Skalek U., Smailus D.E., Schnerch A., Schein J.E.,
RA Jones S.J., Marra M.A.;
RT "Generation and initial analysis of more than 15,000 full-length human
RT and mouse cDNA sequences.";
RL Proc. Natl. Acad. Sci. U.S.A. 99:16899-16903(2002).
RN [2]
RP SEQUENCE FROM N.A.
RC TISSUE=Lung;
RA Strausberg R.;
RL Submitted (JUL-2003) to the EMBL/GenBank/DBJ databases.
DR EMBL; BC054816; AAH54816.1; -.
DR HSSP; P00388; 1J9Z.
DR GO; GO:0010181; F:FMN binding; IEA.
DR GO; GO:0016491; F:oxidoreductase activity; IEA.
DR GO; GO:0006118; P:electron transport; IEA.
DR InterPro; IPR003097; FAD_binding.

```

DR InterPro; IPR008254; Flav nitox synth.
 DR InterPro; IPR001433; Oxred FAD/NAD(P).
 DR Pfam; PF00667; FAD binding_1; 1.
 DR Pfam; PF00258; Flavodoxin_1; 1.
 DR Pfam; PF00175; NAD_binding_1; 1.
 DR PROSITE; PS0902; FLAVODOXIN_LIKE; 1.
 FT NON TER 1
 SQ SEQUENCE 715 AA; 79602 MW; 363A769EA05873CC CRC64;

Query Match 56.4%; Score 44; DB 2; Length 715;
 Best Local Similarity 56.7%; Pred. No. 71;
 Matches 9; Conservative 3; Mismatches 4; Indels 0; Gaps 0;

QY 1 SRLDYKSSLLHLGS 16
 |||:|||||:
 Db 192 SRTDLVASELLHIESQ 207

RESULT 7
 ID Q9A2U5 PRELIMINARY; PRT; 847 AA.
 AC Q9A2U5
 DT 01-JUN-2001 (TrEMBLrel. 17, Created)
 DT 01-JUN-2001 (TrEMBLrel. 17, Last sequence update)
 DT 01-MAR-2004 (TrEMBLrel. 26, Last annotation update)
 DE TonB-dependent receptor.
 GN OrderedLocusNames=CC3461;
 OS Caulobacter crescentus.
 OC Bacteria; Proteobacteria; Alphaproteobacteria; Caulobacteriales;
 OC Caulobacteraceae; Caulobacter.
 OX NCBI_TaxID=155892;
 RN [1]
 RP SEQUENCE FROM N.A.
 RC STRAIN=ATCC 19089 / CB15;
 RX MEDLINE=21173698; PubMed=11259647; DOI=10.1073/pnas.061029298;
 RA Nierman W.C., Feldguy T.V., Laub M.T., Paulsen I.T., Nelson K.E.,
 RA Eisen J.A., Heidelberg J.F., Alley M.R.K., Ohta N., Maddock J.R.,
 RA Potocka I., Nelson W.C., Newton A.S., Stephens C., Phadke N.D., Ely B.,
 RA DeBoy R.T., Dodson R.J., Durkin A.S., Gwinn M.L., Haft D.H.,
 RA Kolonay J.F., Smit J., Craven M.B., Khouri H.M., Shetty J.,
 RA Berry K.J., Utterback T.R., Tran K., Wolf A.M., Vamathevan J.J.,
 RA Ermolaeva M.D., White O., Salzberg S.L., Venter J.C., Shapiro L.,
 RA Fraser C.M.;
 RT "Complete genome sequence of Caulobacter crescentus.";
 RL Proc. Natl Acad. Sci. U.S.A. 98:4136-4141(2001).
 DR EMBL; AE006005; AA25423.1; -.
 DR PIR; C87678; C87678.
 DR TIGR; CC3461; -.
 DR GO; GO:0019867; C:outer membrane; IEA.
 DR GO; GO:0004872; F:receptor activity; IEA.
 DR GO; GO:0005215; F:transporter activity; IEA.
 DR GO; GO:0006810; P:transport; IEA.
 KW Complete proteome; Receptor.
 SQ SEQUENCE 847 AA; 91216 MW; E5882071FBD8344B CRC64;

Query Match 56.4%; Score 44; DB 2; Length 847;
 Best Local Similarity 66.7%; Pred. No. 86;
 Matches 8; Conservative 3; Mismatches 1; Indels 0; Gaps 0;

QY 5 YLKSSLLHLGS 16
 |||:|||||:
 Db 737 YLKASLQHVGNR 748

RESULT 8
 ID Q6LKE8 PRELIMINARY; PRT; 120 AA.
 AC Q6LKE8
 DT 05-JUL-2004 (TrEMBLrel. 27, Created)
 DT 05-JUL-2004 (TrEMBLrel. 27, Last sequence update)
 DT 05-JUL-2004 (TrEMBLrel. 27, Last annotation update)
 DE Hypothetical alkaline phosphatase.
 GN Name=CSAP; OrderedLocusNames=PBPRB0359;

OS Photobacterium profundum (Photobacterium sp. (strain S9)).
 OC Bacteria; Proteobacteria; Gammaproteobacteria; Vibrionales;
 OC Vibrionaceae; Photobacterium.
 OX NCBI_TaxID=74109;
 RN [1]
 RP SEQUENCE FROM N.A.
 RA Vezzi A., Campanaro S., D'Angelo M., Simonato F., Vitulo N., Lauro F.,
 RA Cestaro A., Malacrida G., Simionati B., Cannata N., Bartlett D.,
 RA Valle G.;
 RT "Genome analysis of Photobacterium profundum reveals the complexity of
 RT high pressure adaptations.";
 RL Submitted (MAR-2004) to the EMBL/GenBank/DBJ databases.
 DR EMBL; CR378676; CAG22232.1; -.
 DR HSSP; Q9BHT8; 1K7H.
 KW Complete proteome.
 SQ SEQUENCE 120 AA; 13303 MW; D523E9859B80EE07 CRC64;

Query Match 55.1%; Score 43; DB 2; Length 120;
 Best Local Similarity 61.5%; Pred. No. 15;
 Matches 8; Conservative 3; Mismatches 2; Indels 0; Gaps 0;

QY 3 LDYKSSLLHLGS 15
 :|||: ||: |||
 Db 67 VDYLOQSLRLGS 79

RESULT 9
 ID Q9K8L6 PRELIMINARY; PRT; 216 AA.
 AC Q9K8L6
 DT 01-OCT-2000 (TrEMBLrel. 15, Created)
 DT 01-OCT-2000 (TrEMBLrel. 15, Last sequence update)
 DT 01-JUN-2003 (TrEMBLrel. 24, Last annotation update)
 DE Transcriptional regulator involved in phosphate transport system.
 GN OrderedLocusNames=BH2990;
 OS Bacillus halodurans.
 OC Bacteria; Firmicutes; Bacillales; Bacillaceae; Bacillus.
 OX NCBI_TaxID=86665;
 RN [1]
 RP SEQUENCE FROM N.A.
 RC STRAIN=C-125;
 RX MEDLINE=20512582; PubMed=11058132; DOI=10.1093/nar/28.21.4317;
 RA Takami H., Nakasone K., Takaki Y., Maeno G., Sasaki R., Masui N.,
 RA Fuji F., Hirama C., Nakamura Y., Ogasawara N., Kuhara S.,
 RA Horikoshi K.;
 RT "Complete genome sequence of the alkaliphilic bacterium Bacillus
 RT halodurans and genomic sequence comparison with Bacillus subtilis.";
 RL Nucleic Acids Res. 28:4317-4331(2000).
 DR EMBL; AP001517; BAB06709.1; -.
 DR PIR; F84023; F84023.
 DR InterPro; IPR008170; Phou.
 DR Pfam; PF01895; Phou; 2.
 KW Complete proteome.
 SQ SEQUENCE 216 AA; 24909 MW; D1ABF3CAF66CADC4 CRC64;

Query Match 55.1%; Score 43; DB 2; Length 216;
 Best Local Similarity 57.1%; Pred. No. 29;
 Matches 8; Conservative 3; Mismatches 3; Indels 0; Gaps 0;

QY 1 SRLDYKSSLLHLG 14
 :|||: ||: |||
 Db 8 SLDYLKQGLMMG 21

RESULT 10
 ID Q7MSA0 PRELIMINARY; PRT; 522 AA.
 AC Q7MSA0
 DT 01-MAR-2004 (TrEMBLrel. 26, Created)
 DT 01-MAR-2004 (TrEMBLrel. 26, Last sequence update)
 DT 01-MAR-2004 (TrEMBLrel. 26, Last annotation update)
 DE Hypothetical protein.
 GN OrderedLocusNames=WS0643;

```

OS Wolinella succinogenes.
OC Bacteria; Proteobacteria; Epsilonproteobacteria; Campylobacteriales;
OC Helicobacteraceae; Wolinella.
OX NCBI_TaxID=844;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=DSMZ 1740;
RX MEDLINE=2282897; PubMed=1450908; DOI=10.1073/pnas.1932838100;
RA Baar C., Eppinger M., Raddatz G., Simon J., Lanz C., Klimmek O.,
RA Nandakumar R., Gross R., Rosinus A., Keller H., Jagtap P., Linke B.,
RA Meyer F., Lederer H., Schuster S.C.;
RT "Complete genome sequence and analysis of Wolinella succinogenes.";
RL Proc. Natl. Acad. Sci. U.S.A. 100:11690-11695(2003).
DR EMBL; BX571658; CA509774.1; -.
DR GO; GO:0008484; F:sulfuric ester hydrolase activity; IEA.
DR GO; GO:0008152; P:metabolism; IEA.
DR InterPro; IPR000917; Sulfatase.
DR Pfam; PF00884; Sulfatase; 1.
KW Complete proteome; Hypothetical protein.
SQ SEQUENCE 522 AA; 59741 MW; 999A2343A25C1D62 CRC64;

Query Match 55.1%; Score 43; DB 2; Length 522;
Best Local Similarity 75.0%; Pred. No. 76;
Matches 9; Conservative 1; Mismatches 2; Indels 0; Gaps 0;

QY 4 DYLKSSLLHLGS 15
Db 494 DYLFSSLLHLAN 505

RESULT 11
Q9IN14 PRELIMINARY; PRT; 720 AA.
AC Q9IN14;
DT 01-OCT-2000 (TrEMBLrel. 15, Created)
DT 01-OCT-2000 (TrEMBLrel. 15, Last sequence update)
DT 01-MAR-2001 (TrEMBLrel. 16, Last annotation update)
DE VP3.
OS Banna virus.
OC Viruses; dsRNA viruses; Reoviridae; Seadornavirus.
OX NCBI_TaxID=77763;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=JKT-6423;
RX MEDLINE=20273996; PubMed=10811934;
RA Attoui H., Billoir F., Biagini P., de Micco P., de Lamballerie X.;
RT "Complete sequence determination and genetic analysis of Banna virus
RT and Kadipiro virus: proposal for assignment to a new genus
RT (Seadornavirus) within the family Reoviridae.";
RL J. Gen. Virol. 81:1507-1515(2000).
DR EMBL; AF134515; AAF78856.1; -.
SQ SEQUENCE 720 AA; 82119 MW; 3584BEEA74EDCD66 CRC64;

Query Match 55.1%; Score 43; DB 2; Length 720;
Best Local Similarity 64.3%; Pred. No. 1.1e+02;
Matches 9; Conservative 2; Mismatches 3; Indels 0; Gaps 0;

QY 3 LDYKSSLLHLGSR 16
Db 694 LRVKSKLLHLGSK 707

RESULT 12
OSH3 YEAST STANDARD; PRT; 996 AA.
AC F387I3;
DT 01-FEB-1995 (Rel. 31, Created)
DT 01-FEB-1995 (Rel. 31, Last sequence update)
DT 05-JUL-2004 (Rel. 44, Last annotation update)
DE Oxyesterol-binding protein homolog 3.
GN Name=OSH3; OrderedLocusNames=YHR073W;
OS Saccharomyces cerevisiae (Baker's yeast).
OC Eukaryota; Fungi; Ascomycota; Saccharomycotina; Saccharomycetes;

```

```

OC Saccharomycetales; Saccharomycetaceae; Saccharomyces.
OX NCBI_TaxID=4932;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=S288C / AB972;
RX MEDLINE=94378003; PubMed=8091229;
RA Johnston M., Andrews S., Brinkman R., Cooper J., Ding H., Dover J.,
RA Du Z., Favello A., Fulton L., Gattung S., Geisel C., Kirsten J.,
RA Kucaba T., Hillier L.W., Jier M., Johnston L., Langston Y.,
RA Latreille P., Louis E.J., Macri C., Mardis E., Meneses S., Mouser L.,
RA Nham M., Rifkin L., Riles L., St Peter H., Trevaakis E., Vaughan K.,
RA Vignati D., Wilcox L., Wohlman P., Waterston R., Wilson R.,
RA Vaudin M.;
RT "Complete nucleotide sequence of Saccharomyces cerevisiae chromosome
RT VIII.";
RT Science 265:2077-2082(1994).
RN [2]
RP SUBCELLULAR LOCATION.
RX MEDLINE=21301806; PubMed=11408574;
RA Levine T.P., Munro S.;
RT "Dual targeting of Oshlp, a yeast homologue of oxyesterol-binding
RT protein, to both the Golgi and the nucleus-vacuole junction.";
RL Mol. Biol. Cell 12:1633-1644(2001).
RN [3]
RP GENETIC ANALYSIS.
RX MEDLINE=21135676; PubMed=11238399;
RA Beh C.T., Cool L., Phillips J., Rine J.;
RT "Overlapping functions of the yeast oxyesterol-binding protein
RT homologues.";
RL Genetics 157:1117-1140(2001).
CC -!- SUBCELLULAR LOCATION: Intracellular.
CC -!- SIMILARITY: Belongs to the OSBP family.
CC -!- SIMILARITY: Contains 1 PH domain.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC -----
CC EMBL; U10556; AAB6890.1; -.
CC FIR; S46812; S46812.
CC GerMOnline; 139390; -.
CC SGD; S000001115; OSH3.
CC GO; GO:0005737; C:cytoplasm; IDA.
CC GO; GO:0006694; P:steroid biosynthesis; IGI.
CC InterPro; IPR000648; P:steroid biosynthesis; IGI.
CC InterPro; IPR001849; PH.
CC InterPro; IPR011036; PH related.
CC Pfam; PF01237; Oxyesterol_BP; 1.
CC Pfam; PF00169; PH; 1.
CC SMART; SM00233; PH; 1.
CC PROSITE; PS01013; OSBP; 1.
CC PROSITE; PS00003; PH_DOMAIN; 1.
KW Lipid transport; Transport.
FT DOMAIN 221 315 PH.
SQ SEQUENCE 996 AA; 113759 MW; BIAD1210BFFB9AFE CRC64;

Query Match 55.1%; Score 43; DB 1; Length 996;
Best Local Similarity 69.2%; Pred. No. 1.5e+02;
Matches 9; Conservative 0; Mismatches 4; Indels 0; Gaps 0;

QY 3 LDYKSSLLHLGS 15
Db 64 LDYTKSLHLKGS 76

RESULT 13
Q80UJ4 PRELIMINARY; PRT; 140 AA.
ID Q80UJ4
AC Q80UJ4;

```

DT 01-JUN-2003 (TrEMBLrel. 24, Created)
 DT 01-JUN-2003 (TrEMBLrel. 24, Last sequence update)
 DT 01-MAR-2004 (TrEMBLrel. 26, Last annotation update)
 DE MARVEL (Membrane-associating) domain containing 2.
 GN Name=Marveld2; Synonyms=Mrvldc2;
 OS Mus musculus (Mouse).
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
 OX NCBI_TaxId=10090;
 RN [1]
 RP SEQUENCE FROM N.A.
 RC STRAIN=C2ECH II; TISSUE=Mammary tumor;
 RX MEDLINE=22388257; PubMed=12477932; DOI=10.1073/pnas.242603899;
 RA Strausberg R.L., Feingold E.A., Grouse L.H., Derge J.G.,
 RA Klausner R.D., Collins F.S., Wagner L., Shenmen C.M., Schuler G.D.,
 RA Altschul S.F., Zeeberg B., Buetow K.H., Schaefer C.F., Bhat N.K.,
 RA Hopkins R.F., Jordan H., Moore T., Max S.I., Wang J., Hsieh F.,
 RA Diatchenko L., Marusina K., Farmer A.A., Rubin G.M., Hong L.,
 RA Stapleton M., Soares M.B., Bonaldo M.F., Casavant T.L., Scheetz T.E.,
 RA Brownstein M.J., Uedin T.B., Toshiyuki S., Carninci P., Prange C.,
 RA Raha S.S., Loquellano N.A., Peters G.J., Abramson R.D., Mullaly S.J.,
 RA Bosak S.A., McEwan P.J., McKernan K.J., Malek J.A., Gunaratne P.H.,
 RA Richards S., Worley K.C., Hale S., Garcia A.M., Gay L.J., Hulyk S.W.,
 RA Villalon D.K., Muzny D.M., Sodergren E.J., Lu X., Gibbs R.A.,
 RA Fahey J., Helton E., Kettman M., Madan A., Rodriguez S., Sanchez A.,
 RA Whiting M., Touchman J.W., Green E.D., Dickson M.C.,
 RA Blakeley R.W., Touchman J.W., Green E.D., Dickson M.C.,
 RA Rodriguez A.C., Grimwood J., Schmutz J., Myers R.M., Butterfield Y.S.,
 RA Krzywinski M.I., Skalska U., Smailus D.E., Schnerch A., Schein J.E.,
 RA Jones S.J., Marra M.A.;
 RT "Generation and initial analysis of more than 15,000 full-length human
 and mouse cDNA sequences.";
 RL Proc. Natl. Acad. Sci. U.S.A. 99:16899-16903(2002).
 RP SEQUENCE FROM N.A.
 RC STRAIN=C2ECH II; TISSUE=Mammary tumor;
 RA Strausberg R.;
 RL Submitted (MAR-2003) to the EMBL/GenBank/DBJ databases.
 DR EMBL: BC049919; AAH49919.1; -.
 DR MCD; MGI:2446166; Mrvldc2.
 DR InterPro; IPR010844; Occludin_ELL.
 DR Pfam; PF07303; Occludin_ELL; 1.
 SQ SEQUENCE 140 AA; 16840 MW; 9C7E05E2E6C453F7 CRC64;
 Query Match 53.8%; Score 42; DB 2; Length 140;
 Best Local Similarity 53.3%; Pred. No. 27;
 Matches 8; Conservative 2; Mismatches 5; Indels 0; Gaps 0;
 OY 2 RLDYKSSLLHLGSR 16
 DB 110 RCDYKKNKLSHIKQR 124
 RESULT 14
 Q81YV9 PRELIMINARY; PRT; 350 AA.
 AC Q81YV9;
 DT 01-MAR-2003 (TrEMBLrel. 23, Created)
 DT 01-MAR-2003 (TrEMBLrel. 23, Last sequence update)
 DT 01-MAR-2004 (TrEMBLrel. 26, Last annotation update)
 DE Hypothetical protein MGC34799.
 GN Name=MGC34799;
 OS Homo sapiens (Human).
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 OC Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
 OX NCBI_TaxId=9606;
 RN [1]
 RP SEQUENCE FROM N.A.
 RC TISSUE=Brain;
 RX MEDLINE=22388257; PubMed=12477932; DOI=10.1073/pnas.242603899;
 RA Strausberg R.L., Feingold E.A., Grouse L.H., Derge J.G.,
 RA Klausner R.D., Collins F.S., Wagner L., Shenmen C.M., Schuler G.D.,
 RA Altschul S.F., Zeeberg B., Buetow K.H., Schaefer C.F., Bhat N.K.,

RA Hopkins R.F., Jordan H., Moore T., Max S.I., Wang J., Hsieh F.,
 RA Diatchenko L., Marusina K., Farmer A.A., Rubin G.M., Hong L.,
 RA Stapleton M., Soares M.B., Bonaldo M.F., Casavant T.L., Scheetz T.E.,
 RA Brownstein M.J., Uedin T.B., Toshiyuki S., Carninci P., Prange C.,
 RA Raha S.S., Loquellano N.A., Peters G.J., Abramson R.D., Mullaly S.J.,
 RA Bosak S.A., McEwan P.J., McKernan K.J., Malek J.A., Gunaratne P.H.,
 RA Richards S., Worley K.C., Hale S., Garcia A.M., Gay L.J., Hulyk S.W.,
 RA Villalon D.K., Muzny D.M., Sodergren E.J., Lu X., Gibbs R.A.,
 RA Fahey J., Helton E., Kettman M., Madan A., Rodriguez S., Sanchez A.,
 RA Whiting M., Touchman J.W., Green E.D., Dickson M.C.,
 RA Blakeley R.W., Touchman J.W., Green E.D., Dickson M.C.,
 RA Rodriguez A.C., Grimwood J., Schmutz J., Myers R.M., Butterfield Y.S.,
 RA Krzywinski M.I., Skalska U., Smailus D.E., Schnerch A., Schein J.E.,
 RA Jones S.J., Marra M.A.;
 RT "Generation and initial analysis of more than 15,000 full-length human
 and mouse cDNA sequences.";
 RL Proc. Natl. Acad. Sci. U.S.A. 99:16899-16903(2002).
 RN [2]
 RP SEQUENCE FROM N.A.
 RC TISSUE=Brain;
 RA Strausberg R.;
 RL Submitted (JUL-2002) to the EMBL/GenBank/DBJ databases.
 DR EMBL: BC034769; AAH34769.1; -.
 DR InterPro; IPR003599; IG.
 DR SMART; SM00409; IG; 1.
 KW Hypothetical protein.
 SQ SEQUENCE 350 AA; 38958 MW; 87590C6F93ABCFA9 CRC64;
 Query Match 53.8%; Score 42; DB 2; Length 350;
 Best Local Similarity 61.5%; Pred. No. 74;
 Matches 8; Conservative 3; Mismatches 2; Indels 0; Gaps 0;
 OY 3 LDYKSSLLHLGS 15
 DB 319 IDFIKSSLFGLGS 331
 RESULT 15
 Q99LE8 PRELIMINARY; PRT; 436 AA.
 AC Q99LE8;
 DT 01-JUN-2001 (TrEMBLrel. 17, Created)
 DT 01-JUN-2001 (TrEMBLrel. 17, Last sequence update)
 DT 01-MAR-2004 (TrEMBLrel. 26, Last annotation update)
 DE Mrvldc2 protein (Fragment).
 GN Name=Mrvldc2; Synonyms=BC003296, Mrvldc2;
 OS Mus musculus (Mouse).
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
 OX NCBI_TaxId=10090;
 RN [1]
 RP SEQUENCE FROM N.A.
 RC STRAIN=FVB/N; TISSUE=Mammary tumor;
 RX MEDLINE=22388257; PubMed=12477932; DOI=10.1073/pnas.242603899;
 RA Strausberg R.L., Feingold E.A., Grouse L.H., Derge J.G.,
 RA Klausner R.D., Collins F.S., Wagner L., Shenmen C.M., Schuler G.D.,
 RA Altschul S.F., Zeeberg B., Buetow K.H., Schaefer C.F., Bhat N.K.,
 RA Hopkins R.F., Jordan H., Moore T., Max S.I., Wang J., Hsieh F.,
 RA Diatchenko L., Marusina K., Farmer A.A., Rubin G.M., Hong L.,
 RA Stapleton M., Soares M.B., Bonaldo M.F., Casavant T.L., Scheetz T.E.,
 RA Brownstein M.J., Uedin T.B., Toshiyuki S., Carninci P., Prange C.,
 RA Raha S.S., Loquellano N.A., Peters G.J., Abramson R.D., Mullaly S.J.,
 RA Bosak S.A., McEwan P.J., McKernan K.J., Malek J.A., Gunaratne P.H.,
 RA Richards S., Worley K.C., Hale S., Garcia A.M., Gay L.J., Hulyk S.W.,
 RA Villalon D.K., Muzny D.M., Sodergren E.J., Lu X., Gibbs R.A.,
 RA Fahey J., Helton E., Kettman M., Madan A., Rodriguez S., Sanchez A.,
 RA Whiting M., Touchman J.W., Green E.D., Dickson M.C.,
 RA Blakeley R.W., Touchman J.W., Green E.D., Dickson M.C.,
 RA Rodriguez A.C., Grimwood J., Schmutz J., Myers R.M., Butterfield Y.S.,
 RA Krzywinski M.I., Skalska U., Smailus D.E., Schnerch A., Schein J.E.,
 RA Jones S.J., Marra M.A.;
 RT "Generation and initial analysis of more than 15,000 full-length human
 and mouse cDNA sequences.";

Search completed: November 15, 2005, 12:28:14
Job time : 118.632 secs

This Page Blank (uspto)

GenCore version 5.1.6
Copyright (c) 1993 - 2005 CompuGen Ltd.

OM protein - protein search, using sw model

Run on: November 15, 2005, 12:13:54 ; Search time 32.4211 Seconds
(without alignments)
36.840 Million cell updates/sec

Title: US-10-820-998-1

Perfect score: 78

Sequence: 1 SRLDYKSSLLHLGSR 16

Scoring table: BLOSUM62

Gapop 10.0 , Gapext 0.5

Searched: 513545 seqs, 74649064 residues

Total number of hits satisfying chosen parameters: 513545

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database : Issued Patents AA:*

- 1: /cgn2_6/ptodata/1/iaa/5A_COMB.pep.*
- 2: /cgn2_6/ptodata/1/iaa/5B_COMB.pep.*
- 3: /cgn2_6/ptodata/1/iaa/6A_COMB.pep.*
- 4: /cgn2_6/ptodata/1/iaa/6B_COMB.pep.*
- 5: /cgn2_6/ptodata/1/iaa/PCTUS_COMB.pep.*
- 6: /cgn2_6/ptodata/1/iaa/backfileal.pep.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	49	62.8	129	4	US-09-270-767-32720, A
2	49	62.8	129	4	US-09-270-767-47937, A
3	45	57.7	253	4	US-09-270-767-44896, A
4	44	56.4	724	4	US-09-949-016-10086, A
5	40	51.3	47	4	US-09-702-705-1832, Ap
6	40	51.3	47	4	US-09-736-457-1832, Ap
7	40	51.3	96	4	US-09-702-705-1831, Ap
8	40	51.3	96	4	US-09-736-457-1831, Ap
9	40	51.3	108	4	US-09-702-705-786, Ap
10	40	51.3	108	4	US-09-702-705-793, Ap
11	40	51.3	108	4	US-09-736-457-786, Ap
12	40	51.3	108	4	US-09-736-457-793, Ap
13	40	51.3	108	4	US-09-614-124B-786, Ap
14	40	51.3	108	4	US-09-614-124B-793, Ap
15	40	51.3	108	4	US-09-671-325-786, Ap
16	40	51.3	108	4	US-09-671-325-793, Ap
17	40	51.3	108	4	US-09-589-184-786, Ap
18	40	51.3	108	4	US-09-589-184-793, Ap
19	40	51.3	108	4	US-09-588-824-786, Ap
20	40	51.3	108	4	US-09-588-824-793, Ap
21	40	51.3	126	4	US-09-702-705-791, Ap
22	40	51.3	126	4	US-09-736-457-791, Ap
23	40	51.3	126	4	US-09-614-124B-791, Ap
24	40	51.3	126	4	US-09-671-325-791, Ap
25	40	51.3	126	4	US-09-589-184-791, Ap
26	40	51.3	126	4	US-09-588-824-791, Ap
27	40	51.3	160	4	US-09-702-705-809, Ap

28 40 51.3 160 4 US-09-736-457-809 Sequence 809, App
29 40 51.3 160 4 US-09-614-124B-809 Sequence 809, App
30 40 51.3 160 4 US-09-671-325-809 Sequence 809, App
31 40 51.3 160 4 US-09-589-184-809 Sequence 809, App
32 40 51.3 160 4 US-09-658-824-809 Sequence 809, App
33 40 51.3 192 4 US-09-621-976-5605 Sequence 5605, Ap
34 39 50.0 958 4 US-09-706-594-5 Sequence 5, Appl1
35 38 48.7 73 4 US-09-270-767-62189 Sequence 62189, A
36 38 48.7 131 4 US-09-270-767-39975 Sequence 39975, A
37 38 48.7 131 4 US-09-270-767-55192 Sequence 55192, A
38 38 48.7 200 4 US-09-328-352-7402 Sequence 7402, Ap
39 38 48.7 212 4 US-09-270-767-46591 Sequence 46591, A
40 38 48.7 402 4 US-09-543-681A-5057 Sequence 5057, Ap
41 38 48.7 734 4 US-09-949-016-10597 Sequence 10597, A
42 37 47.4 110 3 US-08-975-762-62 Sequence 62, Appl
43 37 47.4 110 3 US-09-295-028-62 Sequence 62, Appl
44 37 47.4 110 3 US-09-106-582-62 Sequence 62, Appl
45 37 47.4 110 4 US-09-159-469-62 Sequence 62, Appl

ALIGNMENTS

RESULT 1

US-09-270-767-32720
; Sequence 32720, Application US/09270767
; Patent No. 6703491
; GENERAL INFORMATION:
; APPLICANT: Homburger et al.
; TITLE OF INVENTION: Nucleic acids and proteins of Drosophila melanogaster
; FILE REFERENCE: File Reference: 7326-094
; CURRENT APPLICATION NUMBER: US/09/270,767
; CURRENT FILING DATE: 1999-03-17
; NUMBER OF SEQ ID NOS: 62517
; SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 32720
; LENGTH: 129
; TYPE: PRT
; ORGANISM: Drosophila melanogaster
US-09-270-767-32720

Query Match 62.8%; Score 49; DB 4; Length 129;
Best Local Similarity 66.7%; Pred. No. 0.49;
Matches 10; Conservative 1; Mismatches 4; Indels 0; Gaps 0;

QY 1 SRLDYKSSLLHLGSL 15

Db 103 SRFDLLDSQILHLGSL 117

RESULT 2

US-09-270-767-47937
; Sequence 47937, Application US/09270767
; Patent No. 6703491
; GENERAL INFORMATION:
; APPLICANT: Homburger et al.
; TITLE OF INVENTION: Nucleic acids and proteins of Drosophila melanogaster
; FILE REFERENCE: File Reference: 7326-094
; CURRENT APPLICATION NUMBER: US/09/270,767
; CURRENT FILING DATE: 1999-03-17
; NUMBER OF SEQ ID NOS: 62517
; SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 47937
; LENGTH: 129
; TYPE: PRT
; ORGANISM: Drosophila melanogaster
US-09-270-767-47937

Query Match 62.8%; Score 49; DB 4; Length 129;
Best Local Similarity 66.7%; Pred. No. 0.49;
Matches 10; Conservative 1; Mismatches 4; Indels 0; Gaps 0;

QY 1 SRLDYKSSLLHLGSL 15

```

Db      103 SRFDLLSQILHLGS 117
      || | | | : ||||
RESULT 3
US-09-270-767-44896
; Sequence 44896, Application US/09270767
; Patent No. 6703491
; GENERAL INFORMATION:
; APPLICANT: Homburger et al.
; TITLE OF INVENTION: Nucleic acids and proteins of Drosophila melanogaster
; FILE REFERENCE: File Reference: 7326-094
; CURRENT APPLICATION NUMBER: US/09/270,767
; CURRENT FILING DATE: 1999-03-17
; NUMBER OF SEQ ID NOS: 62517
; SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 44896
; LENGTH: 253
; TYPE: PRT
; ORGANISM: Drosophila melanogaster
; FEATURE:
; OTHER INFORMATION: Xaa means any amino acid
US-09-270-767-44896

Query Match      57.7%; Score 45; DB 4; Length 253;
Best Local Similarity 64.3%; Pred. No. 5.3;
Matches 9; Conservative 1; Mismatches 4; Indels 0; Gaps 0;

QY      2 RLDYLKSSLLHLGS 15
      | | | | | | | | :
Db      238 RCQYLKSSLPFLST 251

RESULT 4
US-09-949-016-10086
; Sequence 10086, Application US/09949016
; Patent No. 6812339
; GENERAL INFORMATION:
; APPLICANT: VENTER, J. Craig et al.
; TITLE OF INVENTION: POLYMORPHISMS IN KNOWN GENES ASSOCIATED
; FILE REFERENCE: WITH HUMAN DISEASE, METHODS OF DETECTION AND USES THEREOF
; CURRENT APPLICATION NUMBER: US/09/949,016
; CURRENT FILING DATE: 2000-04-14
; PRIOR APPLICATION NUMBER: 60/241,755
; PRIOR FILING DATE: 2000-10-20
; PRIOR APPLICATION NUMBER: 60/237,768
; PRIOR FILING DATE: 2000-10-03
; PRIOR APPLICATION NUMBER: 60/231,498
; PRIOR FILING DATE: 2000-09-08
; NUMBER OF SEQ ID NOS: 207012
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 10086
; LENGTH: 724
; TYPE: PRT
; ORGANISM: Human
US-09-949-016-10086

Query Match      56.4%; Score 44; DB 4; Length 724;
Best Local Similarity 56.2%; Pred. No. 26;
Matches 9; Conservative 3; Mismatches 4; Indels 0; Gaps 0;

QY      1 SRLDYLKSSLLHLGS 16
      | | | | | | | | :
Db      201 SRTDLVKSELHIESQ 216

RESULT 5
US-09-702-705-1832
; Sequence 1832, Application US/09702705
; Patent No. 6504010
; GENERAL INFORMATION:
; APPLICANT: Wang, Tongtong

```

```

; APPLICANT: Bangur, Chaitanya S.
; APPLICANT: Lodes, Michael A.
; APPLICANT: Fanger, Gary
; APPLICANT: Vedvick, Tom
; APPLICANT: Carter, Darrick
; APPLICANT: Retter, Marc
; APPLICANT: Mannion, Jane
; APPLICANT: Fan, Liqun
; TITLE OF INVENTION: COMPOSITIONS AND METHODS FOR THE THERAPY AND
; FILE REFERENCE: 210121.478C14
; CURRENT APPLICATION NUMBER: US/09/702,705
; CURRENT FILING DATE: 2000-10-30
; NUMBER OF SEQ ID NOS: 1833
; SOFTWARE: FastSeq for Windows Version 3.0
; SEQ ID NO 1832
; LENGTH: 47
; TYPE: PRT
; ORGANISM: Homo sapiens
US-09-702-705-1832

Query Match      51.3%; Score 40; DB 4; Length 47;
Best Local Similarity 72.7%; Pred. No. 5.8;
Matches 8; Conservative 1; Mismatches 2; Indels 0; Gaps 0;

QY      6 LKSSLLHLGSR 16
      | | | | | | | | :
Db      26 LKVGILHLGSR 36

RESULT 6
US-09-736-457-1832
; Sequence 1832, Application US/09736457
; Patent No. 6509448
; GENERAL INFORMATION:
; APPLICANT: Wang, Tongtong
; APPLICANT: Bangur, Chaitanya S.
; APPLICANT: Lodes, Michael A.
; APPLICANT: Fanger, Gary
; APPLICANT: Vedvick, Tom
; APPLICANT: Carter, Darrick
; APPLICANT: Retter, Marc
; APPLICANT: Mannion, Jane
; APPLICANT: Fan, Liqun
; APPLICANT: Wang, Aijun
; TITLE OF INVENTION: COMPOSITIONS AND METHODS FOR THE THERAPY AND
; FILE REFERENCE: 210121.478C15
; CURRENT APPLICATION NUMBER: US/09/736,457
; CURRENT FILING DATE: 2000-12-13
; NUMBER OF SEQ ID NOS: 1864
; SOFTWARE: FastSeq for Windows Version 3.0
; SEQ ID NO 1832
; LENGTH: 47
; TYPE: PRT
; ORGANISM: Homo sapiens
US-09-736-457-1832

Query Match      51.3%; Score 40; DB 4; Length 47;
Best Local Similarity 72.7%; Pred. No. 5.8;
Matches 8; Conservative 1; Mismatches 2; Indels 0; Gaps 0;

QY      6 LKSSLLHLGSR 16
      | | | | | | | | :
Db      26 LKVGILHLGSR 36

RESULT 7
US-09-702-705-1831
; Sequence 1831, Application US/09702705
; Patent No. 6504010
; GENERAL INFORMATION:
; APPLICANT: Wang, Tongtong

```

; APPLICANT: Bangur, Chaitanya S.
 ; APPLICANT: Lodes, Michael A.
 ; APPLICANT: Fanger, Gary
 ; APPLICANT: Vedvick, Tom
 ; APPLICANT: Carter, Darrick
 ; APPLICANT: Retter, Marc
 ; APPLICANT: Mannion, Jane
 ; APPLICANT: Fan, Liqun
 ; TITLE OF INVENTION: COMPOSITIONS AND METHODS FOR THE THERAPY AND
 ; TITLE OF INVENTION: DIAGNOSIS OF LUNG CANCER
 ; FILE REFERENCE: 210121.478C14
 ; CURRENT APPLICATION NUMBER: US/09/702,705
 ; CURRENT FILING DATE: 2000-10-30
 ; NUMBER OF SEQ ID NOS: 1833
 ; SOFTWARE: FastSeq for Windows Version 3.0
 ; SEQ ID NO 1831
 ; LENGTH: 96
 ; TYPE: PRT
 ; ORGANISM: Homo sapiens
 US-09-702-705-1831

Query Match 51.3%; Score 40; DB 4; Length 96;
 Best Local Similarity 72.7%; Pred. No. 13;
 Matches 8; Conservative 1; Mismatches 2; Indels 0; Gaps 0;

QY 6 LKSSLLHLGSR 16
 ||:|||||
 Db 26 LKVGILHLGSR 36

RESULT 8
 US-09-736-457-1831
 ; Sequence 1831, Application US/09736457
 ; Patent No. 6509448
 ; GENERAL INFORMATION:
 ; APPLICANT: Wang, Tongtong
 ; APPLICANT: Bangur, Chaitanya S.
 ; APPLICANT: Lodes, Michael A.
 ; APPLICANT: Fanger, Gary
 ; APPLICANT: Vedvick, Tom
 ; APPLICANT: Carter, Darrick
 ; APPLICANT: Retter, Marc
 ; APPLICANT: Mannion, Jane
 ; APPLICANT: Fan, Liqun
 ; APPLICANT: Wang, Aijun
 ; TITLE OF INVENTION: COMPOSITIONS AND METHODS FOR THE THERAPY AND
 ; TITLE OF INVENTION: DIAGNOSIS OF LUNG CANCER
 ; FILE REFERENCE: 210121.478C15
 ; CURRENT APPLICATION NUMBER: US/09/736,457
 ; CURRENT FILING DATE: 2000-12-13
 ; NUMBER OF SEQ ID NOS: 1864
 ; SOFTWARE: FastSeq for Windows Version 3.0
 ; SEQ ID NO 1831
 ; LENGTH: 96
 ; TYPE: PRT
 ; ORGANISM: Homo sapiens
 US-09-736-457-1831

Query Match 51.3%; Score 40; DB 4; Length 96;
 Best Local Similarity 72.7%; Pred. No. 13;
 Matches 8; Conservative 1; Mismatches 2; Indels 0; Gaps 0;

QY 6 LKSSLLHLGSR 16
 ||:|||||
 Db 26 LKVGILHLGSR 36

RESULT 9
 US-09-702-705-786
 ; Sequence 786, Application US/09702705
 ; Patent No. 6504010
 ; GENERAL INFORMATION:
 ; APPLICANT: Wang, Tongtong

; APPLICANT: Bangur, Chaitanya S.
 ; APPLICANT: Lodes, Michael A.
 ; APPLICANT: Fanger, Gary
 ; APPLICANT: Vedvick, Tom
 ; APPLICANT: Carter, Darrick
 ; APPLICANT: Retter, Marc
 ; APPLICANT: Mannion, Jane
 ; APPLICANT: Fan, Liqun
 ; TITLE OF INVENTION: COMPOSITIONS AND METHODS FOR THE THERAPY AND
 ; TITLE OF INVENTION: DIAGNOSIS OF LUNG CANCER
 ; FILE REFERENCE: 210121.478C14
 ; CURRENT APPLICATION NUMBER: US/09/702,705
 ; CURRENT FILING DATE: 2000-10-30
 ; NUMBER OF SEQ ID NOS: 1833
 ; SOFTWARE: FastSeq for Windows Version 3.0
 ; SEQ ID NO 786
 ; LENGTH: 108
 ; TYPE: PRT
 ; ORGANISM: Homo sapiens
 US-09-702-705-786

Query Match 51.3%; Score 40; DB 4; Length 108;
 Best Local Similarity 72.7%; Pred. No. 15;
 Matches 8; Conservative 1; Mismatches 2; Indels 0; Gaps 0;

QY 6 LKSSLLHLGSR 16
 ||:|||||
 Db 50 LKVGILHLGSR 60

RESULT 10
 US-09-702-705-793
 ; Sequence 793, Application US/09702705
 ; Patent No. 6504010
 ; GENERAL INFORMATION:
 ; APPLICANT: Wang, Tongtong
 ; APPLICANT: Bangur, Chaitanya S.
 ; APPLICANT: Lodes, Michael A.
 ; APPLICANT: Fanger, Gary
 ; APPLICANT: Vedvick, Tom
 ; APPLICANT: Carter, Darrick
 ; APPLICANT: Retter, Marc
 ; APPLICANT: Mannion, Jane
 ; APPLICANT: Fan, Liqun
 ; TITLE OF INVENTION: COMPOSITIONS AND METHODS FOR THE THERAPY AND
 ; TITLE OF INVENTION: DIAGNOSIS OF LUNG CANCER
 ; FILE REFERENCE: 210121.478C14
 ; CURRENT APPLICATION NUMBER: US/09/702,705
 ; CURRENT FILING DATE: 2000-10-30
 ; NUMBER OF SEQ ID NOS: 1833
 ; SOFTWARE: FastSeq for Windows Version 3.0
 ; SEQ ID NO 793
 ; LENGTH: 108
 ; TYPE: PRT
 ; ORGANISM: Homo sapiens
 US-09-702-705-793

Query Match 51.3%; Score 40; DB 4; Length 108;
 Best Local Similarity 72.7%; Pred. No. 15;
 Matches 8; Conservative 1; Mismatches 2; Indels 0; Gaps 0;

QY 6 LKSSLLHLGSR 16
 ||:|||||
 Db 50 LKVGILHLGSR 60

RESULT 11
 US-09-736-457-786
 ; Sequence 786, Application US/09736457
 ; Patent No. 6509448
 ; GENERAL INFORMATION:
 ; APPLICANT: Wang, Tongtong
 ; APPLICANT: Bangur, Chaitanya S.

```

; APPLICANT: Lodes, Michael A.
; APPLICANT: Fanger, Gary
; APPLICANT: Vedvick, Tom
; APPLICANT: Carter, Darrick
; APPLICANT: Retter, Marc
; APPLICANT: Mannion, Jane
; APPLICANT: Fan, Liqun
; APPLICANT: Wang, Aijun
; TITLE OF INVENTION: COMPOSITIONS AND METHODS FOR THE THERAPY AND
; TITLE OF INVENTION: DIAGNOSIS OF LUNG CANCER
; FILE REFERENCE: 210121.478C15
; CURRENT APPLICATION NUMBER: US/09/736,457
; CURRENT FILING DATE: 2000-12-13
; NUMBER OF SEQ ID NOS: 1864
; SOFTWARE: FastSEQ for Windows Version 3.0
; SEQ ID NO 786
; LENGTH: 108
; TYPE: PRT
; ORGANISM: Homo sapiens
US-09-736-457-786

```

```

Query Match      51.3%; Score 40; DB 4; Length 108;
Best Local Similarity 72.7%; Pred. No. 15;
Matches      8; Conservative 1; Mismatches 2; Indels 0; Gaps 0;

```

```

QY      6 LKSSLHLGSR 16
      ||:|||||
Db      50 LKVGILHLGSR 60

```

```

RESULT 12
US-09-736-457-793
; Sequence 793, Application US/09736457
; Patent No. 6509448
; GENERAL INFORMATION:
; APPLICANT: Wang, Tongtong
; APPLICANT: Bangur, Chaitanya S.
; APPLICANT: Lodes, Michael A.
; APPLICANT: Fanger, Gary
; APPLICANT: Vedvick, Tom
; APPLICANT: Carter, Darrick
; APPLICANT: Retter, Marc
; APPLICANT: Mannion, Jane
; APPLICANT: Fan, Liqun
; APPLICANT: Wang, Aijun
; TITLE OF INVENTION: COMPOSITIONS AND METHODS FOR THE THERAPY AND
; TITLE OF INVENTION: DIAGNOSIS OF LUNG CANCER
; FILE REFERENCE: 210121.478C15
; CURRENT APPLICATION NUMBER: US/09/736,457
; CURRENT FILING DATE: 2000-12-13
; NUMBER OF SEQ ID NOS: 1864
; SOFTWARE: FastSEQ for Windows Version 3.0
; SEQ ID NO 793
; LENGTH: 108
; TYPE: PRT
; ORGANISM: Homo sapiens
US-09-736-457-793

```

```

Query Match      51.3%; Score 40; DB 4; Length 108;
Best Local Similarity 72.7%; Pred. No. 15;
Matches      8; Conservative 1; Mismatches 2; Indels 0; Gaps 0;

```

```

QY      6 LKSSLHLGSR 16
      ||:|||||
Db      50 LKVGILHLGSR 60

```

```

RESULT 13
US-09-614-124B-786
; Sequence 786, Application US/09614124B
; Patent No. 6630574
; GENERAL INFORMATION:
; APPLICANT: Wang, Tongtong

```

```

; APPLICANT: Bangur, Chaitanya S.
; APPLICANT: Lodes, Michael A.
; APPLICANT: Fanger, Gary
; APPLICANT: Vedvick, Tom
; APPLICANT: Carter, Darrick
; APPLICANT: Retter, Marc
; APPLICANT: Mannion, Jane
; APPLICANT: Wang, Aijun
; TITLE OF INVENTION: COMPOSITIONS AND METHODS FOR THERAPY AND
; TITLE OF INVENTION: DIAGNOSIS OF LUNG CANCER
; FILE REFERENCE: 210121.478C9
; CURRENT APPLICATION NUMBER: US/09/614,124B
; CURRENT FILING DATE: 2001-07-11
; NUMBER OF SEQ ID NOS: 1668
; SOFTWARE: FastSEQ for Windows Version 3.0
; SEQ ID NO 786
; LENGTH: 108
; TYPE: PRT
; ORGANISM: Homo sapiens
US-09-614-124B-786

```

```

Query Match      51.3%; Score 40; DB 4; Length 108;
Best Local Similarity 72.7%; Pred. No. 15;
Matches      8; Conservative 1; Mismatches 2; Indels 0; Gaps 0;

```

```

QY      6 LKSSLHLGSR 16
      ||:|||||
Db      50 LKVGILHLGSR 60

```

```

RESULT 14
US-09-614-124B-793
; Sequence 793, Application US/09614124B
; Patent No. 6630574
; GENERAL INFORMATION:
; APPLICANT: Wang, Tongtong
; APPLICANT: Bangur, Chaitanya S.
; APPLICANT: Lodes, Michael A.
; APPLICANT: Fanger, Gary
; APPLICANT: Vedvick, Tom
; APPLICANT: Carter, Darrick
; APPLICANT: Retter, Marc
; APPLICANT: Mannion, Jane
; APPLICANT: Wang, Aijun
; TITLE OF INVENTION: COMPOSITIONS AND METHODS FOR THERAPY AND
; TITLE OF INVENTION: DIAGNOSIS OF LUNG CANCER
; FILE REFERENCE: 210121.478C9
; CURRENT APPLICATION NUMBER: US/09/614,124B
; CURRENT FILING DATE: 2001-07-11
; NUMBER OF SEQ ID NOS: 1668
; SOFTWARE: FastSEQ for Windows Version 3.0
; SEQ ID NO 793
; LENGTH: 108
; TYPE: PRT
; ORGANISM: Homo sapiens
US-09-614-124B-793

```

```

Query Match      51.3%; Score 40; DB 4; Length 108;
Best Local Similarity 72.7%; Pred. No. 15;
Matches      8; Conservative 1; Mismatches 2; Indels 0; Gaps 0;

```

```

QY      6 LKSSLHLGSR 16
      ||:|||||
Db      50 LKVGILHLGSR 60

```

```

RESULT 15
US-09-671-325-786
; Sequence 786, Application US/09671325
; Patent No. 6667154
; GENERAL INFORMATION:
; APPLICANT: Wang, Tongtong
; APPLICANT: Bangur, Chaitanya S.
; APPLICANT: Lodes, Michael A.
; APPLICANT: Fanger, Gary

```

; APPLICANT: Vedvick, Tom
; APPLICANT: Carter, Darrick
; APPLICANT: Retter, Marc
; APPLICANT: Mannion, Jane
; APPLICANT: Fan, Liqun
; TITLE OF INVENTION: COMPOSITIONS AND METHODS FOR THE THERAPY AND
; FILE REFERENCE: 210121.478C12
; CURRENT APPLICATION NUMBER: US/09/671,325
; CURRENT FILING DATE: 2000-09-26
; NUMBER OF SEQ ID NOS: 1825
; SOFTWARE: FastSeq for Windows Version 3.0
; SEQ ID NO 786
; LENGTH: 108
; TYPE: PRT
; ORGANISM: Homo sapiens
US-09-671-325-786

Query Match 51.3%; Score 40; DB 4; Length 108;
Best Local Similarity 72.7%; Pred. No. 15;
Matches 8; Conservative 1; Mismatches 2; Indels 0; Gaps 0;

Qy 6 LKSSLLHLGSR 16
||:|||||
Db 50 LKVGILHLGSR 60

Search completed: November 15, 2005, 12:29:35
Job time : 33.4211 secs

This Page Blank (uspto)

GenCore version 5.1.6
Copyright (c) 1993 - 2005 Compugen Ltd.

OM protein - protein search, using sw model

Run on: November 15, 2005, 12:12:49 ; Search time 80.1842 Seconds
(without alignments)
70.249 Million cell updates/sec

Title: US-10-820-998-3

Perfect score: 59

Sequence: 1 CCCSSSSSSSS 11

Scoring table: BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 1612378 seqs, 512079187 residues

Total number of hits satisfying chosen parameters: 1612378

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database : Uniprot 03.*

1: uniprot_sprot.*

2: uniprot_trembl.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	56	94.9	268	2	O22798 arabidopsis
2	52	88.1	55	1	P41927 yarrowia li
3	52	88.1	55	1	MT1_YARLI
4	50	84.7	47	2	O9hfd0 yarrowia li
5	50	84.7	107	2	O76nt4 dictyosteli
6	50	84.7	613	2	O9d673 mus musculu
7	50	84.7	799	2	O94087 agastache r
8	50	84.7	989	2	O61587 oryza sativ
9	50	84.7	1031	2	O43302 homo sapien
10	47	79.7	157	2	O9h462 dictyosteli
11	47	79.7	593	1	O9u944 schizosacch
12	47	79.7	620	2	O42709 schizosacch
13	47	79.7	727	2	O9fuw5 schizosacch
14	47	79.7	727	2	O8w364 oryza sativ
15	47	79.7	727	2	O7xd99 oryza sativ
16	47	79.7	885	2	O61232 lymanaea sta
17	47	79.7	1307	2	O8gv37 arabidopsis
18	47	79.7	1307	2	O98710 arabidopsis
19	46	78.0	86	2	O70b05 ipomoea yel
20	46	78.0	165	2	O6lu37 photobacter
21	46	78.0	471	2	O8lsy2 arabidopsis
22	46	78.0	639	2	O9ln44 arabidopsis
23	46	78.0	1922	2	O7p2e8 anopheles g
24	45	76.3	113	2	O9v841 drosophila
25	45	76.3	161	2	O9skb5 arabidopsis
26	45	76.3	568	2	O6ill1 drosophila
27	45	76.3	608	2	O8rwr8 zea mays (m
28	45	76.3	877	2	O8s7t9 oryza sativ
29	44	74.6	99	2	O7xl03 oryza sativ
30	44	74.6	99	2	O7vq88 chlamydia p
31	44	74.6	110	2	O9k218 chlamydia p
					O8vav8 white spot

RESULT 1

ID	O22798	PRELIMINARY;	PRT;	268 AA.
AC	O22798;			
DT	01-JAN-1998 (TrEMBLrel. 05, Created)			
DT	01-JAN-1998 (TrEMBLrel. 05, Last sequence update)			
DT	05-JUL-2004 (TrEMBLrel. 27, Last annotation update)			
DE	Putative NAM (No apical meristem)-like protein			
DE	(At2g33480/F4P9.25).			
GN	Name=At2g33480;			
OS	Arabidopsis thaliana (Mouse-ear cress).			
OC	Eukaryota; Viridiplantae; Streptophyta; Tracheophyta;			
OC	Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots; rosids;			
OC	eurosid II; Brassicales; Brassicaceae; Arabidopsi			
OX	NCBI_TaxID=3702;			
RN	[1]			
RP	SEQUENCE FROM N.A.			
RA	Rounsley S.D., Lin X., Ketchum K.A., Crosby M.L., Brandon R.C., C.R.,			
RA	Sykes S.M., Mason T.M., Kerlavage A.R., Adams M.D., Somerville			
RA	Venter J.C.;			
RL	Submitted (MAR-2000) to the EMBL/GenBank/DBJ databases.			
RN	[2]			
RP	SEQUENCE FROM N.A.			
RA	Town C.D., Kaul S.;			
RL	Submitted (FEB-2002) to the EMBL/GenBank/DBJ databases.			
RN	[3]			
RP	SEQUENCE FROM N.A.			
RA	Shinn P., Chen H., Cheuk R., Kim C.J., Meyers M.C., Banh J.,			
RA	Bowser L., Carninci P., Chang E., Dale J.M., Goldsmith A.D.,			
RA	Hayashizaki Y., Ishida J., Jones T., Kamiya A., Karlin-Neumann G.,			
RA	Kawai J., Lam B., Lee J.M., Lin J., Miranda M., Narusaka M.,			
RA	Nguyen M., Onodera C.S., Palm C.J., Quach H.L., Sakurai T., Satou M.,			
RA	Seki M., Southwick A., Tang C.C., Toriumi M., Wu H.C., Yamada K.,			
RA	Yamamura Y., Yu G., Yu S., Shinozaki K., Davis R.W., Theologis A.,			
RA	Ecker J.R.;			
RL	Submitted (MAR-2002) to the EMBL/GenBank/DBJ databases.			
RN	[4]			
RP	SEQUENCE FROM N.A.			
RA	Southwick A., Karlin-Neumann G., Nguyen M., Lam B., Miranda M.,			
RA	Palm C.J., Theologis A., Ecker J., Davis R.W.;			
RL	Submitted (NOV-2000) to the EMBL/GenBank/DBJ databases.			
RN	[5]			
RP	SEQUENCE FROM N.A.			
RA	Cheuk R., Chen H., Kim C.J., Koesema E., Meyers M.C., Banh J.,			
RA	Bowser L., Carninci P., Dale J.M., Goldsmith A.D., Hayashizaki Y.,			
RA	Ishida J., Jiang P.X., Jones T., Kamiya A., Karlin-Neumann G.,			
RA	Kawai J., Lam B., Lee J.M., Lin J., Liu S.X., Miranda M., Narusaka M.,			
RA	Nguyen M., Onodera C.S., Palm C.J., Pham P.K., Quach H.L., Sakurai T.,			
RA	Satou M., Seki M., Southwick A., Tang C.C., Toriumi M., Yamada K.,			
RA	Yamamura Y., Yu G., Yu S., Shinozaki K., Davis R.W., Theologis A.,			
RA	Ecker J.R.;			
RL	Submitted (AUG-2001) to the EMBL/GenBank/DBJ databases.			
DR	EMBL; AC002332; AAB80665.1; -			
DR	EMBL; AY093730; AAW10354.1; -			

32	44	74.6	239	2	O96AJ6	O96AJ6 homo sapien
33	44	74.6	334	2	O7SGF2	O7SGF2 neurospora
34	44	74.6	340	2	O6WVF9	O6WVF9 neurospora
35	44	74.6	365	2	O7QI98	O7QI98 anopheles g
36	44	74.6	417	2	O9XYQ7	O9XYQ7 lytechinus
37	44	74.6	722	2	O7S563	O7S563 neurospora
38	44	74.6	1233	2	O7PSV6	O7PSV6 anopheles g
39	44	74.6	1820	1	STR9_HUMAN	STR9p6 homo sapien
40	43	72.9	79	1	LC81_ARATH	LC81v0 arabidopsis
41	43	72.9	186	2	O627V0	O627V0 oryza sativ
42	43	72.9	199	2	O8JNA5	O8JNA5 fringilla c
43	43	72.9	294	2	O9XUS0	O9XUS0 caenorhabdi
44	43	72.9	415	1	SCR_DROME	P09077 drosophila
45	43	72.9	517	2	O6K8X7	O6K8X7 oryza sativ

ALIGNMENTS

Yarrowia lipolytica (Candida lipolytica); Eukaryota; Fungi; Ascomycota; Saccharomycotina; Saccharomycetes; Yarrowia.

Q9D673 PRELIMINARY; PRT; 107 AA.
AC Q9D673;
DT 01-JUN-2001 (TrEMBLrel. 17, Created)
DT 01-JUN-2001 (TrEMBLrel. 17, Last sequence update)
DT 01-OCT-2002 (TrEMBLrel. 22, Last annotation update)
DE Mus musculus 0 day neonate skin cDNA, RIKEN full-length enriched
DE library, clone:4632413K17 product:hypothetical protein, full insert
DE sequence.
GN Name=4632413K17Rik;
OS Mus musculus (Mouse).
OC Eukaryota; Metazoa; Chordata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
OX NCBI_TaxID=10090;
RN [1]
RN SEQUENCE FROM N.A.
RC STRAIN=C57BL/6J; TISSUE=Skin;
RC STRAIN=C57BL/6J; TISSUE=Skin;
RX MEDLINE=9279253; PubMed=10349636; DOI=10.1016/S0076-6879(99)03004-9;
RA Carninci P., Hayashizaki Y.;
RT "High-efficiency full-length cDNA cloning.";
RL Meth. Enzymol. 303:19-44(1999).
RN [2]
RN SEQUENCE FROM N.A.
RC STRAIN=C57BL/6J; TISSUE=Skin;
RX MEDLINE=21085650; PubMed=11217851; DOI=10.1038/35055500;
RA RIKEN FANTOM Consortium;
RT "Functional annotation of a full-length mouse cDNA collection.";
RL Nature 409:685-690(2001).
RN [3]
RN SEQUENCE FROM N.A.
RC STRAIN=C57BL/6J; TISSUE=Skin;
RC The FANTOM Consortium;
RA The RIKEN Genome Exploration Research Group Phase I & II Team;
RT "Analysis of the mouse transcriptome based on functional annotation of
RT 60,770 full-length cDNAs.";
RL Nature 420:563-573(2002).
RN [4]
RN SEQUENCE FROM N.A.
RC STRAIN=C57BL/6J; TISSUE=Skin;
RX MEDLINE=20499374; PubMed=11042159; DOI=10.1101/gr.145100;
RA Carninci P., Shibata Y., Hayatsu N., Sugahara Y., Shibata K., Itoh M.,
RA Konno H., Okazaki Y., Muramatsu M., Hayashizaki Y.;
RT "Normalization and subtraction of cap-trapper-selected cDNAs to
RT prepare full-length cDNA libraries for rapid discovery of new genes.";
RL Genome Res. 10:1617-1630(2000).
RN [5]
RN SEQUENCE FROM N.A.
RC STRAIN=C57BL/6J; TISSUE=Skin;
RX MEDLINE=20530913; PubMed=11076861; DOI=10.1101/gr.152600;
RA Shibata K., Itoh M., Aizawa K., Nagaoka S., Sasaki N., Carninci P.,
RA Konno H., Akiyama J., Nishi K., Kitsuunai T., Tashiro H., Itoh M.,
RA Sumi N., Ishii Y., Nakamura S., Hazama M., Nishine T., Haxada A.,
RA Yamanoto R., Matsumoto H., Sakaguchi S., Ikegami T., Kashiwagi K.,
RA Fujiwaki S., Inoue K., Togawa Y., Izawa M., Ohara E., Watahiki M.,
RA Yoneda Y., Ishikawa T., Ozawa K., Tanaka T., Matsuura S., Kawai J.,
RA Okazaki Y., Muramatsu M., Inoue Y., Kira A., Hayashizaki Y.;
RT "RIKEN integrated sequence analysis (RISA) system-384-format
RT sequencing pipeline with 384 multicapillary sequencer.";
RL Genome Res. 10:1757-1771(2000).
RN [6]
RN SEQUENCE FROM N.A.
RC STRAIN=C57BL/6J; TISSUE=Skin;
RA Adachi J., Aizawa K., Akahira S., Akimura T., Atai A., Aono H.,
RA Arakawa T., Bono H., Carninci P., Fukuda S., Fukunishi Y., Furuno M.,
RA Hanagaki T., Hara A., Hayatsu N., Hiramoto K., Hiraoka T., Hori F.,
RA Imotani K., Ishii Y., Itoh M., Izawa M., Kasukawa T., Kato H.,
RA Kawai J., Kojima Y., Konno H., Kouda M., Koya S., Kurihara C.,
RA Matsuyama T., Miyazaki A., Nishi K., Nomura K., Numazaki R., Ohno M.,
RA Okazaki Y., Okido T., Owa C., Saito H., Saito R., Sakai C., Sakai K.,
RA Sano H., Saeki D., Shibata K., Shibata Y., Shinagawa A., Shiraki T.,
RA Sogabe Y., Suzuki H., Tagami M., Tagawa A., Takahashi F., Tanaka T.,
RA Tejima Y., Toya T., Yamamura T., Yasunishi A., Yoshida K., Yoshino M.,
RA Muramatsu M., Hayashizaki Y.;

RL Submitted (JUL-2000) to the EMBL/GenBank/DBJ databases.
DR EMBL; AK014574; BAB29438.1; -;
DR MGD; MGI:1924301; 4632413K17Rik.
KW Hypothetical protein.
SQ SEQUENCE 107 AA; 11062 MW; 38A570E3683A86FC CRC64;
Query Match 84.7%; Score 50; DB 2; Length 107;
Best Local Similarity 100.0%; Pred. No. 9.2;
Matches 10; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 2 CCSSSSSSSS 11
DB 38 CCSSSSSSSS 47
RESULT 6
Q940E7 PRELIMINARY; PRT; 613 AA.
AC Q940E7;
DT 01-DEC-2001 (TrEMBLrel. 19, Created)
DT 01-DEC-2001 (TrEMBLrel. 19, Last sequence update)
DT 01-OCT-2003 (TrEMBLrel. 25, Last annotation update)
DE D-limonene synthase.
OS Agastache rugosa.
OC Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
OC Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots; asterids;
OC Lamiales; Lamiaceae; Nepetoideae; Nepeteae; Agastache.
OX NCBI_TaxID=39271;
RN [1]
RN SEQUENCE FROM N.A.
RA Saeki D., Maruyama T., Ito M., Honda G.;
RL Submitted (SEP-2001) to the EMBL/GenBank/DBJ databases.
DR EMBL; AY055214; AAL17636.1; -;
DR HSSP; O81192; 1N1Z.
DR GO; GO:0016829; F-lyase activity; IEA.
DR GO; GO:0008152; P-metabolism; IEA.
DR InterPro; IPR005630; Terpene synthase.
DR InterPro; IPR008949; Terpenoid synthase.
DR InterPro; IPR008930; Terpene synthase.
DR InterPro; IPR001906; Terpene synthase.
DR Pfam; PF01397; Terpene synthase 1.
DR Pfam; PF03936; Terpene synthase C; 1.
SQ SEQUENCE 613 AA; 71422 MW; A92502C399F47572 CRC64;
Query Match 84.7%; Score 50; DB 2; Length 613;
Best Local Similarity 100.0%; Pred. No. 49;
Matches 10; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 2 CCSSSSSSSS 11
DB 54 CCSSSSSSSS 63
RESULT 7
Q615E7 PRELIMINARY; PRT; 799 AA.
AC Q615E7;
DT 05-JUL-2004 (TrEMBLrel. 27, Created)
DT 05-JUL-2004 (TrEMBLrel. 27, Last sequence update)
DT 05-JUL-2004 (TrEMBLrel. 27, Last annotation update)
DE Hypothetical protein OSUNBA002J22.1;
GN Name=OSUNBA002J22.1;
OS Oryza sativa (japonica cultivar-group).
OC Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
OC Spermatophyta; Magnoliophyta; Liliopsida; Poales; Poaceae;
OC Ehrhartoideae; Oryzaceae; Oryza.
OX NCBI_TaxID=39947;
RN [1]
RN SEQUENCE FROM N.A.
RA Chow T.-Y., Heing Y.-I.C., Chen C.-S., Chen H.-H., Liu S.-M.,
RA Chao Y.-T., Chang S.-J., Chen H.-C., Chen S.-K., Chen T.-R.,
RA Chen Y.-L., Cheng C.-H., Chung C.-I., Han S.-Y., Hsiao S.-H.,
RA Hsiung J.-N., Hsu C.-H., Huang J.-J., Kau P.-I., Lee M.-C., Leu H.-L.,

```

RA Li Y.-F., Lin S.-J., Lin Y.-C., Wu S.-W., Yu C.-Y., Yu S.-W.,
RA Wu H.-P., Shaw J.-F.;
RL Submitted (JUN-2004) to the EMBL/GenBank/DBJ databases.
DR EMBL: AC136216; AAT58808.1; -.
DR GO: GO:0042025; C:host cell nucleus; IEA.
DR GO: GO:0016987; F:sigma factor activity; IEA.
DR GO: GO:0003700; F:transcription factor activity; IEA.
DR GO: GO:0006260; P:DNA replication; IEA.
DR GO: GO:0006355; P:regulation of transcription, DNA-dependent; IEA.
DR GO: GO:0006350; P:transcription; IEA.
DR GO: GO:0006352; P:transcription initiation; IEA.
DR InterPro: IPR007627; Sigma70_r2.
DR InterPro: IPR007624; Sigma70_r3.
DR InterPro: IPR007630; Sigma70_r4.
DR InterPro: IPR000943; Sigma70.
DR InterPro: IPR003176; Vir DNA binding.
DR Pfam: PF04542; Sigma70_r2; 1.
DR Pfam: PF04539; Sigma70_r3; 1.
DR Pfam: PF04545; Sigma70_r4; 1.
DR PRINTS: PR00046; SIGMA70FCT.
KW Hypothetical protein.
SQ SEQUENCE 799 AA; 90227 MW; 23B49D4B0E93590D CRC64;

Query Match 84.7%; Score 50; DB 2; Length 799;
Best Local Similarity 100.0%; Pred. No. 63;
Matches 10; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Oy 2 CCSSSSSSSS 11
Db 256 CCSSSSSSSS 265

RESULT 8
O43302 PRELIMINARY; PRT; 989 AA.
ID O43302
AC O43302
DT 01-JUN-1998 (TrEMBLrel. 06, Created)
DT 01-MAR-2004 (TrEMBLrel. 26, Last sequence update)
DT 01-MAR-2004 (TrEMBLrel. 26, Last annotation update)
DE KIAA0418 protein (Fragment).
GN Name=KIAA0418;
OS Homo sapiens (Human).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
OX NCBI_TaxID=9606;
RN [1]
RP SEQUENCE FROM N.A.
RC TISSUE=Brain;
RX MEDLINE=98116655; PubMed=9455477;
RA Iehikawa K., Nagase T., Nakajima D., Seki N., Ohira M., Miyajima N.,
RA Tanaka A., Kotani H., Nomura N., Ohara O.;
RT "Prediction of the coding sequences of unidentified human genes. VIII.
RT 78 new cDNA clones from brain which code for large proteins in
RT vitro.";
RL DNA Res. 4:307-313(1997).
CC -!- SIMILARITY: Contains 5 SH3 domains.
DR EMBL: AB007878; BAA24848.2; -.
DR PIR: T00056; T00056.
DR HSP; P14598; 10V3.
DR InterPro: IPR001452; SH3.
DR Pfam: PF00018; SH3.1; 5.
DR ProDom: PD000066; SH3; 5.
DR SMART: SM00326; SH3; 5.
DR PROSITE; PS50002; SH3; 5.
KW SH3 domain.
FT NON TER 1
SQ SEQUENCE 989 AA; 107451 MW; A41E53FE65EE916B CRC64;

Query Match 84.7%; Score 50; DB 2; Length 989;
Best Local Similarity 100.0%; Pred. No. 77;
Matches 10; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Oy 2 CCSSSSSSSS 11

```

```

Db 556 CCSSSSSSSS 565
|||||||
RESULT 9
Q9H462 PRELIMINARY; PRT; 1031 AA.
ID Q9H462
AC Q9H462
DT 01-MAR-2001 (TrEMBLrel. 16, Created)
DT 01-MAR-2001 (TrEMBLrel. 16, Last sequence update)
DT 01-MAR-2003 (TrEMBLrel. 23, Last annotation update)
DE BA16N2.2 (Similar to murine FISH (An SH3 and PX domain-containing
DE protein, and Src substructure)) (Fragment).
GN Name=BA16N2.2;
OS Homo sapiens (Human).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
OX NCBI_TaxID=9606;
RN [1]
RP SEQUENCE FROM N.A.
RA Chapman J.;
RL Submitted (AUG-2000) to the EMBL/GenBank/DBJ databases.
CC -!- SIMILARITY: Contains 5 SH3 domains.
DR EMBL: ALI21929; CAC10184.1; -.
DR HSP; P14598; 10V3.
DR Genew; HGNC:23664; SH3MD1.
DR InterPro: IPR001452; SH3.
DR Pfam: PF00018; SH3.1; 5.
DR ProDom: PD000066; SH3; 5.
DR SMART; SM00326; SH3; 5.
DR PROSITE; PS50002; SH3; 5.
KW SH3 domain.
FT NON TER 1
SQ SEQUENCE 1031 AA; 113126 MW; CF673CB96D2443E7 CRC64;

Query Match 84.7%; Score 50; DB 2; Length 1031;
Best Local Similarity 100.0%; Pred. No. 80;
Matches 10; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Oy 2 CCSSSSSSSS 11
Db 598 CCSSSSSSSS 607

RESULT 10
Q9U9A4 PRELIMINARY; PRT; 157 AA.
ID Q9U9A4
AC Q9U9A4
DT 01-MAY-2000 (TrEMBLrel. 13, Created)
DT 01-MAY-2000 (TrEMBLrel. 13, Last sequence update)
DT 01-MAR-2004 (TrEMBLrel. 26, Last annotation update)
DE SAPLP C protein.
OS Dictyostelium discoideum (Slime mold).
OC Eukaryota; Mycetozoa; Dictyostelidia; Dictyostelium.
OX NCBI_TaxID=44689;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=AX2;
RA Rudolph M., Nickel R., Leippe M.;
RL Submitted (AUG-1999) to the EMBL/GenBank/DBJ databases.
DR EMBL; AF175768; AAD49964.1; -.
DR DictyBase; DDB0191411; sala.
DR GO: GO:0006629; P:lipid metabolism; IEA.
DR InterPro: IPR007856; SapB_1.
DR InterPro: IPR008138; SapB_2.
DR InterPro: IPR008139; SaposinB.
DR InterPro: IPR011001; Saposin_like.
DR Pfam; PF05184; SapB_1; 1.
DR Pfam; PF03489; SapB_2; 1.
DR SMART; SM00741; SapB; 1.
SQ SEQUENCE 157 AA; 16684 MW; D10007A01B1E4FA2 CRC64;

Query Match 79.7%; Score 47; DB 2; Length 157;

```

```
Best Local Similarity 90.9%; Pred. No. 32;
Matches 10; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 1 CCCCCSSSSSS 11
Db 99 CCDSSSSSSSSS 109

RESULT 11
MCMA_SCHPO STANDARD; PRT; 593 AA.
ID AC02709;
DT 16-OCT-2001 (Rel. 40, Created)
DT 16-OCT-2001 (Rel. 40, Last sequence update)
DT 05-JUL-2004 (Rel. 44, Last annotation update)
DE DNA replication licensing factor mcm10 (Minichromosome maintenance
DE protein 10) (Cdc23 protein).
GN Names:mcm10; Synonyms:cdc23; ORFNames:SPBC1347.10;
OS Schizosaccharomyces pombe (fission yeast).
OC Eukaryota; Fungi; Ascomycota; Schizosaccharomycetes;
OC Schizosaccharomycetales; Schizosaccharomycetaceae;
OC Schizosaccharomycetes.
OX NCBI_TaxID=4896;
RN [1]
RP SEQUENCE FROM N.A., AND FUNCTION.
RX MEDLINE=98441643; PubMed=9745018;
RA Aves S.J., Tongue N., Foster A.J., Hart E.A.;
RT "The essential Schizosaccharomycetes pombe cdc23 DNA replication gene
RT shares structural and functional homology with the Saccharomycetes
RT cerevisiae DNA43 (MCM10) gene.";
RL Curr. Genet. 34:164-171(1998).
RN [2]
RP SEQUENCE FROM N.A., FUNCTION, SUBUNIT, AND SUBCELLULAR LOCATION.
RC STRAIN=972;
RX MEDLINE=20245491; PubMed=10783164;
RA Homesley L., Lei M., Kawasaki Y., Sawyer S., Christensen T., Tye B.K.;
RT "Mcm10 and the MCM2-7 complex interact to initiate DNA synthesis and
RT to release replication factors from origins.";
RL Genes Dev. 14:913-926(2000).
RN [3]
RP SEQUENCE FROM N.A.
RC STRAIN=972;
RX MEDLINE=21848401; PubMed=11859360; DOI=10.1038/nature724;
RA Wood V., Williams R., Rajandream M.A., Lyne M., Lyne R., Stewart A.,
RA Sgouros J., Peat N., Hayles J., Baker S., Basham D., Bowman S.,
RA Brooks K., Brown D., Brown S., Chillingworth T., Churcher C.M.,
RA Collins M., Connor R., Cronin A., Davis P., Feltwell T., Fraser A.,
RA Gentles S., Goble A., Hamlin N., Harris D., Hidalgo J., Hodgson G.,
RA Holroyd S., Hornsby T., Howarth S., Huckle E.J., Hunt S., Jagels K.,
RA James K., Jones L., Jones M., Leather S., McDonald S., McLean J.,
RA Mooney P., Moule S., Mungall K., Murphy L., Niblett D., Odell C.,
RA Oliver K., O'Neil S., Pearson D., Quail M.A., Rabinowitsch E.,
RA Rutherford K., Rutter S., Saunders D., Seeger K., Sharp S.,
RA Skellton J., Simmonds M., Squares R., Squares S., Stevens K.,
RA Taylor K., Taylor R.G., Tivey A., Walsh S.V., Warren T., Whitehead S.,
RA Woodward J., Volkhardt G., Aert R., Robben J., Grymonprez B.,
RA Weltjens I., Vanstreels E., Rieger M., Schaefer M., Mueller-Auer S.,
RA Gabel C., Fuchs M., Fritzc C., Holzer E., Mosel D., Hilbert H.,
RA Borzym K., Langer I., Beck A., Lehrach H., Reinhardt R., Pohl T.M.,
RA Eger P., Zimmermann W., Wedler H., Wambutt R., Fumelle B.,
RA Goffeau A., Cadieu E., Dreano S., Gloux S., Lelaure V., Mottier S.,
RA Galibert F., Aves S.J., Xiang Z., Hunt C., Moore K., Hurst S.M.,
RA Lucas M., Rochet M., Gaillardin C., Tallada V.A., Garzon A., Thode G.,
RA Daga R.R., Cruzado L., Jimenez J., Sanchez M., del Rey F., Benito J.,
RA Dominguez A., Revuelta J.L., Moreno S., Armstrong J., Forsburg S.L.,
RA Cerutti L., Lowe T., McCombie W.R., Paulsen I., Potashkin J.,
RA Shpakovskii G.V., Ussery D., Barrell B.G., Nurse P.;
RT "The genome sequence of Schizosaccharomycetes pombe.";
RL Nature 415:871-880(2002).
CC -!- FUNCTION: Required for DNA synthesis. Required for entry into or
CC completion of S phase.
CC -!- SUBUNIT: Interacts with mcm7.
CC -!- SUBCELLULAR LOCATION: Nuclear. Associates with chromatin.
```

```
CC -!- SIMILARITY: To yeast DNA43.
CC This SWISS-PROT entry is copyrighted. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC -----
DR EMBL; AJ224944; CAAL2235.1; -
DR EMBL; AB011244; BAA24935.1; -
DR EMBL; AL035548; CAB37441.1; -
DR FIR; T43323; T43323.
DR GenDB; Spombe; SPBC1347.10; -
DR InterPro; IPR004365; tRNA_anti_1.
DR Pfam; PF01336; tRNA_anti_1.
KW Cell cycle; Cell division; Nuclear protein.
FT DOMAIN 422 429
SQ SEQUENCE 593 AA; 66633 MW; AA2CDB29E74BD349 CRC64;

Query Match 79.7%; Score 47; DB 1; Length 593;
Best Local Similarity 81.8%; Pred. No. 1.2e+02;
Matches 9; Conservative 1; Mismatches 1; Indels 0; Gaps 0;

QY 1 CCCCCSSSSSS 11
Db 419 CCASSSSSSSNS 429

RESULT 12
Q9F0W5 PRELIMINARY; PRT; 620 AA.
AC Q9F0W5;
DT 01-WAR-2001 (TrEMBLrel. 16, Created)
DT 01-OCT-2001 (TrEMBLrel. 18, Last sequence update)
DT 01-OCT-2003 (TrEMBLrel. 25, Last annotation update)
DE (+)-4R-limonene synthase.
OS Schizonepeta tenuifolia.
OC Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots; asterids;
OC Lamiales; Lamiaceae; Nepetoideae; Nepeteae; Schizonepeta.
OX NCBI_TaxID=135200;
RN [1]
RP SEQUENCE FROM N.A.
RX MEDLINE=21200190; PubMed=11305598;
RA Maruyama T., Ito M., Kiuchi F., Honda G.;
RT "Molecular cloning, functional expression and characterization of d-
RT limonene synthase from Schizonepeta tenuifolia.";
RL Biol. Pharm. Bull. 24:373-377(2001).
DR EMBL; AF282875; AAG01140.2; -
DR HSP; O81192; 1N12.
DR GO; GO:0016829; F-lyase activity; IEA.
DR GO; GO:0008152; P-metabolism; IEA.
DR InterPro; IPR005630; Terpene synth. C.
DR InterPro; IPR008949; Terpenoid synth.
DR InterPro; IPR008930; Terp cyc toroid.
DR InterPro; IPR001906; Terp synth-like.
DR Pfam; PF01397; Terpene synth; 1.
DR Pfam; PF03936; Terpene synth; C. 1.
SQ SEQUENCE 620 AA; 71809 MW; 9A8E7302442C55E CRC64;

Query Match 79.7%; Score 47; DB 2; Length 620;
Best Local Similarity 90.0%; Pred. No. 1.2e+02;
Matches 9; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

QY 2 CCCCCSSSSSS 11
Db 57 CCASSSSSSS 66

RESULT 13
Q8W364
```

```

ID Q8W364 PRELIMINARY; PRT; 727 AA.
AC Q8W364;
DT 01-MAR-2002 (TrEMBLrel. 20, Created)
DT 01-MAR-2002 (TrEMBLrel. 20, Last sequence update)
DT 01-OCT-2003 (TrEMBLrel. 25, Last annotation update)
DE Hypochemical protein OSUNB0029C15.6;
GN Name=OSUNB0029C15.6;
OS Oryza sativa (Rice).
OC Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
OC Spermatophyta; Magnoliophyta; Liliopsida; Poales; Poaceae;
OC Ehrhartoideae; Oryzeae; Oryza.
OX NCBI_TaxID=4530;
RN [1]
RP SEQUENCE FROM N.A.
RA Buell C.R., Yuan Q., Ouyang S., Liu J., Moffat K.S., Hill J.N.,
RA Gansberger K., Brenner M., Burgess S., Hance M., Shvartsbeyn M.,
RA Tsitrin T., Riggs F., Hsiao J., Zismann V., Blunt S., Pai G.,
RA vanAken S.E., Uterback T.R., Feldblyum T.V., Kalb E., Quackenbush J.,
RA Salzberg S.L., White O., Fraser C.M.;
RL Submitted (DEC-2000) to the EMBL/GenBank/DBJ databases.
RN [2]
RP SEQUENCE FROM N.A.
RA Buell R.;
RL Submitted (JAN-2002) to the EMBL/GenBank/DBJ databases.
DR EMBL; AC087182; AAL59023.1; -.
DR Gramene; Q8W364; -.
KW Hypochemical protein.
SQ SEQUENCE 727 AA; 81207 MW; FFA15A1786CFCACE CRC64;

Query Match 79.7%; Score 47; DB 2; Length 727;
Best Local Similarity 90.0%; Pred. No. 1.4e+02;
Matches 9; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

Qy 2 CCSSSSSSSS 11
|||||:|
Db 22 CCSSSSSSSS 31

RESULT 14
Q7XD99 PRELIMINARY; PRT; 727 AA.
AC Q7XD99;
DT 01-OCT-2003 (TrEMBLrel. 25, Created)
DT 01-OCT-2003 (TrEMBLrel. 25, Last sequence update)
DT 01-OCT-2003 (TrEMBLrel. 25, Last annotation update)
DE Hypochemical protein.
GN ORFNames=OSUNB0029C15.6;
OS Oryza sativa (japonica cultivar-group).
OC Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
OC Spermatophyta; Magnoliophyta; Liliopsida; Poales; Poaceae;
OC Ehrhartoideae; Oryzeae; Oryza.
OX NCBI_TaxID=39947;
RN [1]
RP SEQUENCE FROM N.A.
RA The Rice Chromosome 10 Sequencing Consortium;
RA "In-depth view of structure, activity, and evolution of rice
RT chromosome 10.";
RL Science 300:1566-1569(2003).
RN [2]
RP SEQUENCE FROM N.A.
RA Buell C.R., Wing R.A., McCombie W.R., Messing J., Yuan Q.;
RL Submitted (MAY-2003) to the EMBL/GenBank/DBJ databases.
DR EMBL; AE017107; AAP54343.1; -.
DR Gramene; Q7XD99; -.
KW Hypochemical protein.
SQ SEQUENCE 727 AA; 81207 MW; FFA15A1786CFCACE CRC64;

Query Match 79.7%; Score 47; DB 2; Length 727;
Best Local Similarity 90.0%; Pred. No. 1.4e+02;
Matches 9; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

Qy 2 CCSSSSSSSS 11
|||||:|

```

```

Db 22 CCSSSSSSSS 31

RESULT 15
O61232 PRELIMINARY; PRT; 885 AA.
ID O61232;
AC O61232;
DT 01-AUG-1998 (TrEMBLrel. 07, Created)
DT 01-AUG-1998 (TrEMBLrel. 07, Last sequence update)
DT 01-MAR-2004 (TrEMBLrel. 26, Last annotation update)
DE Serotonin receptor 5-HT2.
OS Lymnaea stagnalis (Great pond snail).
OC Eukaryota; Metazoa; Mollusca; Gastropoda; Pulmonata; Basommatophora;
OC Lymnaeidae; Lymnaeidae; Lymnaea.
OX NCBI_TaxID=6523;
RN [1]
RP SEQUENCE FROM N.A.
RX MEDLINE=97046685; PubMed=8891606; DOI=10.1016/0014-2999(96)00410-4;
RA Gerhardt C.C., Leyssen J.E., Planta R.J., Vreugdenhil E.,
RA Van Heerikhuizen H.;
RT "Functional characterisation of a 5-HT2 receptor cDNA cloned from
RT Lymnaea stagnalis.";
RL Eur. J. Pharmacol. 311:249-258(1996).
DR EMBL; U50080; AAC16969.1; -.
DR GO; GO:0016021; C:integral to membrane; IEA.
DR GO; GO:0004872; F:receptor activity; IEA.
DR GO; GO:0001584; F:rhodopsin-like receptor activity; IEA.
DR GO; GO:0007186; P:G-protein coupled receptor protein signalin. .; IEA.
DR InterPro; IPR000276; GPCR_Rhodopsn.
DR Pfam; PF00001; 7tm_1; 1.
DR PRINTS; PR00237; GPCRHHODOPSN.
DR PROSITE; PS00237; G_PROTEIN_RECEP_F1_1; UNKNOWN_1.
DR PROSITE; PS02462; G_PROTEIN_RECEP_F1_2; 1.
KW Receptor.
SQ SEQUENCE 885 AA; 97577 MW; 40E0CAA2105B5216 CRC64;

Query Match 79.7%; Score 47; DB 2; Length 885;
Best Local Similarity 90.0%; Pred. No. 1.7e+02;
Matches 9; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

Qy 2 CCSSSSSSSS 11
|||||:|
Db 686 CCTSSSSSSS 695

Search completed: November 15, 2005, 12:28:16
Job time : 82.1842 sec

```

GenCore version 5.1.6
Copyright (c) 1993 - 2005 Compugen Ltd.

OM protein - protein search, using sw model

Run on: November 15, 2005, 12:13:54 ; Search time 22.2895 Seconds
(without alignments)
36.840 Million cell updates/sec

Title: US-10-820-998-3
Perfect score: 59
Sequence: 1 CCCSSSSSSSS 11

Scoring table: BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 513545 seqs, 74649064 residues

Total number of hits satisfying chosen parameters: 513545

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : Issued Patents AA: *
1: /cgn2_6/ptodata/1/iaa/5A_COMB.pep.*
2: /cgn2_6/ptodata/1/iaa/5B_COMB.pep.*
3: /cgn2_6/ptodata/1/iaa/6A_COMB.pep.*
4: /cgn2_6/ptodata/1/iaa/6B_COMB.pep.*
5: /cgn2_6/ptodata/1/iaa/PCTUS_COMB.pep.*
6: /cgn2_6/ptodata/1/iaa/backfile1.pep.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	49	83.1	125	4	US-09-248-796A-27793
2	46.5	78.8	438	4	US-09-248-796A-23919
3	46	78.0	143	4	US-09-252-991A-18285
4	45	76.3	284	4	US-09-248-796A-25423
5	44	74.6	104	4	US-09-198-452A-117
6	43	72.9	99	4	US-09-205-258-1133
7	43	72.9	187	4	US-09-252-991A-20667
8	43	72.9	280	4	US-09-252-991A-17338
9	43	72.9	594	4	US-09-949-016-9764
10	43	72.9	596	4	US-09-949-016-6303
11	43	72.9	858	1	PCT-US92-04227-2
12	43	72.9	858	5	PCT-US92-04227-2
13	42	71.2	70	4	US-09-248-796A-25823
14	42	71.2	148	4	US-09-270-767-39762
15	42	71.2	148	4	US-09-270-767-54979
16	42	71.2	214	3	US-09-199-637A-35
17	42	71.2	245	4	US-09-252-991A-30321
18	42	71.2	869	4	US-09-252-991A-16746
19	41	69.5	34	4	US-09-270-767-40588
20	41	69.5	34	4	US-09-270-767-55804
21	41	69.5	47	4	US-09-471-276-1133
22	41	69.5	137	4	US-09-252-991A-31370
23	41	69.5	140	4	US-09-252-991A-30876
24	41	69.5	163	2	US-08-727-688-25
25	41	69.5	175	4	US-10-101-464A-690
26	41	69.5	218	4	US-09-252-991A-21934
27	41	69.5	256	4	US-09-270-767-37173

28	41	69.5	256	4	US-09-270-767-52390	Sequence 52390, A
29	41	69.5	305	4	US-09-252-991A-27334	Sequence 27334, A
30	41	69.5	379	4	US-09-252-991A-28428	Sequence 28428, A
31	41	69.5	493	4	US-08-979-847B-106	Sequence 106, App
32	41	69.5	609	4	US-09-538-092-711	Sequence 711, App
33	41	69.5	632	4	US-09-270-767-42178	Sequence 42178, A
34	41	69.5	737	4	US-09-866-028-15	Sequence 15, Appl
35	41	69.5	737	4	US-09-944-457-15	Sequence 15, Appl
36	41	69.5	970	4	US-09-270-767-42741	Sequence 42741, A
37	40	67.8	67	4	US-09-270-767-34335	Sequence 34335, A
38	40	67.8	67	4	US-09-270-767-49552	Sequence 49552, A
39	40	67.8	75	3	US-08-927-219-51	Sequence 51, Appl
40	40	67.8	119	4	US-09-252-991A-33059	Sequence 33059, A
41	40	67.8	133	4	US-09-252-991A-25367	Sequence 25367, A
42	40	67.8	170	4	US-09-252-991A-22362	Sequence 22362, A
43	40	67.8	183	4	US-09-252-991A-17497	Sequence 17497, A
44	40	67.8	246	3	US-09-185-160-7	Sequence 7, Appl
45	40	67.8	272	4	US-09-252-991A-31152	Sequence 31152, A

ALIGNMENTS

RESULT 1
US-09-248-796A-27793
; Sequence 27793, Application US/09248796A
; Patent No. 6747137
; GENERAL INFORMATION:
; APPLICANT: Keith Weinstock et al
; TITLE OF INVENTION: NUCLEIC ACID AND AMINO ACID SEQUENCES RELATING TO CANDIDA ALBICANS
; FILE REFERENCE: 107196.132
; CURRENT APPLICATION NUMBER: US/09/248,796A
; CURRENT FILING DATE: 1999-02-12
; PRIOR APPLICATION NUMBER: US 60/074,725
; PRIOR FILING DATE: 1998-02-13
; PRIOR APPLICATION NUMBER: US 60/096,409
; PRIOR FILING DATE: 1998-08-13
; NUMBER OF SEQ ID NOS: 28208
; SEQ ID NO 27793
; LENGTH: 125
; TYPE: PRT
; ORGANISM: Candida albicans
US-09-248-796A-27793

Query Match 83.1%; Score 49; DB 4; Length 125;
Best Local Similarity 90.0%; Pred. No. 24;
Matches 9; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 1 CCCSSSSSS 10
||| |||||
Db 6 CCCSSSSSS 15

RESULT 2
US-09-248-796A-23919
; Sequence 23919, Application US/09248796A
; Patent No. 6747137
; GENERAL INFORMATION:
; APPLICANT: Keith Weinstock et al
; TITLE OF INVENTION: NUCLEIC ACID AND AMINO ACID SEQUENCES RELATING TO CANDIDA ALBICANS
; FILE REFERENCE: 107196.132
; CURRENT APPLICATION NUMBER: US/09/248,796A
; CURRENT FILING DATE: 1999-02-12
; PRIOR APPLICATION NUMBER: US 60/074,725
; PRIOR FILING DATE: 1998-02-13
; PRIOR APPLICATION NUMBER: US 60/096,409
; PRIOR FILING DATE: 1998-08-13
; NUMBER OF SEQ ID NOS: 28208
; SEQ ID NO 23919
; LENGTH: 438
; TYPE: PRT

; ORGANISM: Candida albicans
; FEATURE:
; NAME/KEY: UNSURE
; LOCATION: (3), (4)
; OTHER INFORMATION: Identity of amino acid sequences at the above locations are unknown
US-09-248-796A-23919

Query Match 78.8%; Score 46.5; DB 4; Length 438;
Best Local Similarity 68.8%; Pred. No. 1.4e+02;
Matches 11; Conservative 0; Mismatches 0; Indels 5; Gaps 1;

Qy 1 CCCSSSS-----SSSS 11
| | | | | | | | | | |
Db 256 CCCSSSSCCYCCSSSS 271

RESULT 3
US-09-252-991A-18285
; Sequence 18285, Application US/09252991A
; Patent No. 6551795
; GENERAL INFORMATION:
; APPLICANT: Marc J. Rubenfield et al.
; TITLE OF INVENTION: NUCLEIC ACID AND AMINO ACID SEQUENCES RELATING TO PSEUDOMONAS
; FILE OF INVENTION: AERUGINOSA FOR DIAGNOSTICS AND THERAPEUTICS
; FILE REFERENCE: 107196.136
; CURRENT APPLICATION NUMBER: US/09/252.991A
; CURRENT FILING DATE: 1999-02-18
; PRIOR APPLICATION NUMBER: US 60/074,788
; PRIOR FILING DATE: 1998-02-18
; PRIOR APPLICATION NUMBER: US 60/094,190
; PRIOR FILING DATE: 1998-07-27
; NUMBER OF SEQ ID NOS: 33142
; SEQ ID NO 18285
; LENGTH: 143
; TYPE: PRT
; ORGANISM: Pseudomonas aeruginosa
US-09-252-991A-18285

Query Match 78.0%; Score 46; DB 4; Length 143;
Best Local Similarity 72.7%; Pred. No. 59;
Matches 8; Conservative 1; Mismatches 2; Indels 0; Gaps 0;

Qy 1 CCCSSSSSSSS 11
| | | | | | | | | | |
Db 11 CCCSSSACRSSS 21

RESULT 4
US-09-248-796A-25423
; Sequence 25423, Application US/09248796A
; Patent No. 6747137
; GENERAL INFORMATION:
; APPLICANT: Keith Weinstein et al
; TITLE OF INVENTION: NUCLEIC ACID AND AMINO ACID SEQUENCES RELATING TO CANDIDA ALBICAN
; FILE OF INVENTION: FOR DIAGNOSTICS AND THERAPEUTICS
; FILE REFERENCE: 107196.132
; CURRENT APPLICATION NUMBER: US/09/248,796A
; CURRENT FILING DATE: 1999-02-12
; PRIOR APPLICATION NUMBER: US 60/074,725
; PRIOR FILING DATE: 1998-02-13
; PRIOR APPLICATION NUMBER: US 60/096,409
; PRIOR FILING DATE: 1998-08-13
; NUMBER OF SEQ ID NOS: 28208
; SEQ ID NO 25423
; LENGTH: 284
; TYPE: PRT
; ORGANISM: Candida albicans
; FEATURE:
; NAME/KEY: UNSURE
; LOCATION: (250), (279)
; OTHER INFORMATION: Identity of amino acid sequences at the above locations are unknown
US-09-248-796A-25423

Query Match 76.3%; Score 45; DB 4; Length 284;
Best Local Similarity 90.0%; Pred. No. 1.4e+02;
Matches 9; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

Qy 1 CCCSSSSSSSS 10
| | | | | | | | | | |
Db 98 CSCSSSSSSSS 107

RESULT 5
US-09-198-452A-117
; Sequence 117, Application US/09198452A
; Patent No. 6559294
; GENERAL INFORMATION:
; APPLICANT: Griflais, R.
; TITLE OF INVENTION: Chlamydia pneumoniae genomic sequence and polypeptides, fragments
; TITLE OF INVENTION: thereof and uses thereof, in particular for the diagnosis, prevention
; FILE OF INVENTION: and treatment of infection
; FILE REFERENCE: 9710-003-999
; CURRENT APPLICATION NUMBER: US/09/198,452A
; CURRENT FILING DATE: 1998-11-24
; NUMBER OF SEQ ID NOS: 6849
; SEQ ID NO 117
; LENGTH: 104
; TYPE: PRT
; ORGANISM: Chlamydia pneumoniae
US-09-198-452A-117

Query Match 74.8%; Score 44; DB 4; Length 104;
Best Local Similarity 63.6%; Pred. No. 76;
Matches 7; Conservative 1; Mismatches 3; Indels 0; Gaps 0;

Qy 1 CCCSSSSSSSS 11
| | | | | | | | | | |
Db 26 CCCSDSSSKA 36

RESULT 6
US-09-205-258-1133
; Sequence 1133, Application US/09205258
; Patent No. 6525174
; GENERAL INFORMATION:
; APPLICANT: Young et al.
; TITLE OF INVENTION: 207 Human Secreted Proteins
; FILE REFERENCE: P2007P1
; CURRENT APPLICATION NUMBER: US/09/205,258
; CURRENT FILING DATE: 1998-12-04
; EARLIER APPLICATION NUMBER: PCT/US98/11422
; EARLIER FILING DATE: 1998-06-04
; EARLIER APPLICATION NUMBER: 60/048,885
; EARLIER FILING DATE: 1997-06-06
; EARLIER APPLICATION NUMBER: 60/049,375
; EARLIER FILING DATE: 1997-06-06
; EARLIER APPLICATION NUMBER: 60/048,881
; EARLIER FILING DATE: 1997-06-06
; EARLIER APPLICATION NUMBER: 60/048,880
; EARLIER FILING DATE: 1997-06-06
; EARLIER APPLICATION NUMBER: 60/048,896
; EARLIER FILING DATE: 1997-06-06
; EARLIER APPLICATION NUMBER: 60/049,020
; EARLIER FILING DATE: 1997-06-06
; EARLIER APPLICATION NUMBER: 60/048,876
; EARLIER FILING DATE: 1997-06-06
; EARLIER APPLICATION NUMBER: 60/048,895
; EARLIER FILING DATE: 1997-06-06
; EARLIER APPLICATION NUMBER: 60/048,884
; EARLIER FILING DATE: 1997-06-06
; EARLIER APPLICATION NUMBER: 60/048,894
; EARLIER FILING DATE: 1997-06-06
; EARLIER APPLICATION NUMBER: 60/048,971
; EARLIER FILING DATE: 1997-06-06
; EARLIER APPLICATION NUMBER: 60/048,964
; EARLIER FILING DATE: 1997-06-06

EARLIER APPLICATION NUMBER: 60/048,882
EARLIER FILING DATE: 1997-06-06
EARLIER APPLICATION NUMBER: 60/048,899
EARLIER FILING DATE: 1997-06-06
EARLIER APPLICATION NUMBER: 60/048,893
EARLIER FILING DATE: 1997-06-06
EARLIER APPLICATION NUMBER: 60/048,900
EARLIER FILING DATE: 1997-06-06
EARLIER APPLICATION NUMBER: 60/048,901
EARLIER FILING DATE: 1997-06-06
EARLIER APPLICATION NUMBER: 60/048,892
EARLIER FILING DATE: 1997-06-06
EARLIER APPLICATION NUMBER: 60/048,915
EARLIER FILING DATE: 1997-06-06
EARLIER APPLICATION NUMBER: 60/048,919
EARLIER FILING DATE: 1997-06-06
EARLIER APPLICATION NUMBER: 60/048,970
EARLIER FILING DATE: 1997-06-06
EARLIER APPLICATION NUMBER: 60/048,972
EARLIER FILING DATE: 1997-06-06
EARLIER APPLICATION NUMBER: 60/048,916
EARLIER FILING DATE: 1997-06-06
EARLIER APPLICATION NUMBER: 60/049,373
EARLIER FILING DATE: 1997-06-06
EARLIER APPLICATION NUMBER: 60/048,917
EARLIER FILING DATE: 1997-06-06
EARLIER APPLICATION NUMBER: 60/048,883
EARLIER FILING DATE: 1997-06-06
EARLIER APPLICATION NUMBER: 60/048,897
EARLIER FILING DATE: 1997-06-06
EARLIER APPLICATION NUMBER: 60/048,898
EARLIER FILING DATE: 1997-06-06
EARLIER APPLICATION NUMBER: 60/048,962
EARLIER FILING DATE: 1997-06-06
EARLIER APPLICATION NUMBER: 60/048,963
EARLIER FILING DATE: 1997-06-06
EARLIER APPLICATION NUMBER: 60/048,877
EARLIER FILING DATE: 1997-06-06
EARLIER APPLICATION NUMBER: 60/048,878
EARLIER FILING DATE: 1997-06-06
EARLIER APPLICATION NUMBER: 60/070,923
EARLIER FILING DATE: 1997-12-18
EARLIER APPLICATION NUMBER: 60/092,921
EARLIER FILING DATE: 1998-07-15
EARLIER APPLICATION NUMBER: 60/094,657
EARLIER FILING DATE: 1998-07-30
NUMBER OF SEQ ID NOS: 1227
SOFTWARE: PatentIn Ver. 2.0
SEQ ID NO 1133
LENGTH: 99
TYPE: PRT
ORGANISM: Homo sapiens
US-09-252-991A-20667

Query Match 72.9%; Score 43; DB 4; Length 99;
Best Local Similarity 80.0%; Pred. No. 95;
Matches 8; Conservative 1; Mismatches 1; Indels 0; Gaps 0;

Qy 2 CCCCCSSSSSS 11
| | | | | | | | | |
Db 66 CCCCCSSSSSS 75

RESULT 7
US-09-252-991A-20667

Sequence 20667, Application US/09252991A
Patent No. 6551795
GENERAL INFORMATION:
APPLICANT: Marc J. Rubenfield et al.
TITLE OF INVENTION: NUCLEIC ACID AND AMINO ACID SEQUENCES RELATING TO PSEUDOMONAS
FILE REFERENCE: 107196.136
CURRENT APPLICATION NUMBER: US/09/252,991A
CURRENT FILING DATE: 1999-02-18
PRIOR APPLICATION NUMBER: US 60/074,788
PRIOR FILING DATE: 1998-02-18
PRIOR APPLICATION NUMBER: US 60/094,190
PRIOR FILING DATE: 1998-07-27
NUMBER OF SEQ ID NOS: 33142
SEQ ID NO 20667
LENGTH: 187
TYPE: PRT
ORGANISM: Pseudomonas aeruginosa
US-09-252-991A-20667

Query Match 72.9%; Score 43; DB 4; Length 187;
Best Local Similarity 54.5%; Pred. No. 1.6e+02;
Matches 6; Conservative 4; Mismatches 1; Indels 0; Gaps 0;

Qy 1 CCCCCSSSSSS 11
| | | | | | | | | |
Db 119 CCGGTAATSA 129

RESULT 8

US-09-252-991A-17338
Sequence 17338, Application US/09252991A
Patent No. 6551795
GENERAL INFORMATION:
APPLICANT: Marc J. Rubenfield et al.
TITLE OF INVENTION: NUCLEIC ACID AND AMINO ACID SEQUENCES RELATING TO PSEUDOMONAS
FILE REFERENCE: 107196.136
CURRENT APPLICATION NUMBER: US/09/252,991A
CURRENT FILING DATE: 1999-02-18
PRIOR APPLICATION NUMBER: US 60/074,788
PRIOR FILING DATE: 1998-02-18
PRIOR APPLICATION NUMBER: US 60/094,190
PRIOR FILING DATE: 1998-07-27
NUMBER OF SEQ ID NOS: 33142
SEQ ID NO 17338
LENGTH: 280
TYPE: PRT
ORGANISM: Pseudomonas aeruginosa
US-09-252-991A-17338

Query Match 72.9%; Score 43; DB 4; Length 280;
Best Local Similarity 72.7%; Pred. No. 2.3e+02;
Matches 8; Conservative 2; Mismatches 1; Indels 0; Gaps 0;

Qy 1 CCCCCSSSSSS 11
| | | | | | | | | |
Db 46 CCCCCSSSSSSA 56

RESULT 9

US-09-949-016-9764
Sequence 9764, Application US/09949016
Patent No. 6812339
GENERAL INFORMATION:
APPLICANT: VENTER, J. Craig et al.
TITLE OF INVENTION: POLYMORPHISMS IN KNOWN GENES ASSOCIATED
FILE REFERENCE: CL001307
CURRENT APPLICATION NUMBER: US/09/949,016
CURRENT FILING DATE: 2000-04-14
PRIOR APPLICATION NUMBER: 60/241,755
PRIOR FILING DATE: 2000-10-20

; PRIOR APPLICATION NUMBER: 60/237,768
 ; PRIOR FILING DATE: 2000-10-03
 ; PRIOR APPLICATION NUMBER: 60/231,498
 ; PRIOR FILING DATE: 2000-09-08
 ; NUMBER OF SEQ ID NOS: 207012
 ; SOFTWARE: FastSeq for Windows Version 4.0
 ; SEQ ID NO 9764
 ; LENGTH: 594
 ; TYPE: PRT
 ; ORGANISM: Human
 US-09-949-016-9764

Query Match 72.9%; Score 43; DB 4; Length 594;
 Best Local Similarity 58.8%; Pred. No. 4.4e+02;
 Matches 10; Conservative 1; Mismatches 0; Indels 6; Gaps 1;

Qy 1 CCCSS-----SSSS 11
 Db 224 CCCASSGAGAAASSSS 240

RESULT 10

US-09-949-016-6303
 ; Sequence 6303, Application US/09949016
 ; Patent No. 6812339
 ; GENERAL INFORMATION:
 ; APPLICANT: VENTER, J. Craig et al.
 ; TITLE OF INVENTION: POLYMORPHISMS IN KNOWN GENES ASSOCIATED
 ; WITH HUMAN DISEASE, METHODS OF DETECTION AND USES THEREOF
 ; FILE REFERENCE: CL001307
 ; CURRENT APPLICATION NUMBER: US/09/949,016
 ; CURRENT FILING DATE: 2000-04-14
 ; PRIOR APPLICATION NUMBER: 60/241,755
 ; PRIOR FILING DATE: 2000-10-20
 ; PRIOR APPLICATION NUMBER: 60/237,768
 ; PRIOR FILING DATE: 2000-10-03
 ; PRIOR APPLICATION NUMBER: 60/231,498
 ; PRIOR FILING DATE: 2000-09-08
 ; NUMBER OF SEQ ID NOS: 207012
 ; SOFTWARE: FastSeq for Windows Version 4.0
 ; SEQ ID NO 6303
 ; LENGTH: 596
 ; TYPE: PRT
 ; ORGANISM: Human
 US-09-949-016-6303

Query Match 72.9%; Score 43; DB 4; Length 596;
 Best Local Similarity 58.8%; Pred. No. 4.4e+02;
 Matches 10; Conservative 1; Mismatches 0; Indels 6; Gaps 1;

Qy 1 CCCSS-----SSSS 11
 Db 226 CCCASSGAGAAASSSS 242

RESULT 11

US-07-712-284-2
 ; Sequence 2, Application US/07712284
 ; Patent No. 5227158
 ; GENERAL INFORMATION:
 ; APPLICANT: Jardieu, Paula M.
 ; TITLE OF INVENTION: Hepatocyte Growth Stimulation
 ; NUMBER OF SEQUENCES: 2
 ; CORRESPONDENCE ADDRESS:
 ; ADDRESSEE: Genentech, Inc.
 ; STREET: 460 Point San Bruno Blvd
 ; CITY: South San Francisco
 ; STATE: California
 ; COUNTRY: USA
 ; ZIP: 94080
 ; COMPUTER READABLE FORM:
 ; MEDIUM TYPE: 5.25 inch, 360 Kb floppy disk
 ; COMPUTER: IBM PC compatible

; OPERATING SYSTEM: PC-DOS/MS-DOS
 ; SOFTWARE: patin (Genentech)
 ; CURRENT APPLICATION DATA:
 ; APPLICATION NUMBER: US/07/712,284
 ; FILING DATE: 19910610
 ; CLASSIFICATION: 424
 ; PRIOR APPLICATION DATA:
 ; APPLICATION NUMBER:
 ; FILING DATE:
 ; ATTORNEY/AGENT INFORMATION:
 ; NAME: Dreger, Ginger R.
 ; REGISTRATION NUMBER: 33,055
 ; REFERENCE/DOCKET NUMBER: 704
 ; TELECOMMUNICATION INFORMATION:
 ; TELEPHONE: 415/266-3216
 ; TELEFAX: 415/952-9881
 ; TELEX: 910/371-7168
 ; INFORMATION FOR SEQ ID NO: 2:
 ; SEQUENCE CHARACTERISTICS:
 ; LENGTH: 858 amino acids
 ; TYPE: AMINO ACID
 ; TOPOLOGY: linear
 US-07-712-284-2

Query Match 72.9%; Score 43; DB 1; Length 858;
 Best Local Similarity 72.7%; Pred. No. 6e+02;
 Matches 8; Conservative 0; Mismatches 3; Indels 0; Gaps 0;

Qy 1 CCCSSSSSSSS 11
 Db 44 CCCSSSSCISS 54

RESULT 12

PCT-US92-04227-2
 ; Sequence 2, Application PC/TUS9204227
 ; GENERAL INFORMATION:
 ; APPLICANT: GENENTECH, INC.
 ; APPLICANT: Jardieu, Paula M.
 ; TITLE OF INVENTION: Hepatocyte Growth Stimulation
 ; NUMBER OF SEQUENCES: 2
 ; CORRESPONDENCE ADDRESS:
 ; ADDRESSEE: Genentech, Inc.
 ; STREET: 460 Point San Bruno Blvd
 ; CITY: South San Francisco
 ; STATE: California
 ; COUNTRY: USA
 ; ZIP: 94080
 ; COMPUTER READABLE FORM:
 ; MEDIUM TYPE: 5.25 inch, 360 Kb floppy disk
 ; COMPUTER: IBM PC compatible
 ; OPERATING SYSTEM: PC-DOS/MS-DOS
 ; SOFTWARE: patin (Genentech)
 ; CURRENT APPLICATION DATA:
 ; APPLICATION NUMBER: PCT/US92/04227
 ; FILING DATE: 19920519
 ; CLASSIFICATION: 424
 ; PRIOR APPLICATION DATA:
 ; APPLICATION NUMBER: 07/712,284
 ; FILING DATE: 10 June 1991
 ; ATTORNEY/AGENT INFORMATION:
 ; NAME: Dreger, Ginger R.
 ; REGISTRATION NUMBER: 33,055
 ; REFERENCE/DOCKET NUMBER: 704P1
 ; TELECOMMUNICATION INFORMATION:
 ; TELEPHONE: 415/225-3216
 ; TELEFAX: 415/952-9881
 ; TELEX: 910/371-7168
 ; INFORMATION FOR SEQ ID NO: 2:
 ; SEQUENCE CHARACTERISTICS:
 ; LENGTH: 858 amino acids
 ; TYPE: AMINO ACID
 ; TOPOLOGY: linear

PCT-US92-04227-2

Query Match 72.9%; Score 43; DB 5; Length 858;
Best Local Similarity 72.7%; Pred. No. 6e+02;
Matches 8; Conservative 0; Mismatches 3; Indels 0; Gaps 0;

Qy 1 CCCSSSSSSSS 11
Db 44 CCCSMSSCISS 54

RESULT 13

US-09-248-796A-25823
; Sequence 25823, Application US/09248796A
; Patent No. 6747137

; GENERAL INFORMATION:

; APPLICANT: Keith Weinstock et al

; TITLE OF INVENTION: NUCLEIC ACID AND AMINO ACID SEQUENCES RELATING TO CANDIDA ALBICANS

; FILE REFERENCE: 107196.132

; CURRENT APPLICATION NUMBER: US/09/248,796A

; CURRENT FILING DATE: 1999-02-12

; PRIOR APPLICATION NUMBER: US 60/074,725

; PRIOR FILING DATE: 1998-02-13

; PRIOR APPLICATION NUMBER: US 60/096,409

; PRIOR FILING DATE: 1998-08-13

; NUMBER OF SEQ ID NOS: 28208

; SEQ ID NO 25823

; LENGTH: 70

; TYPE: PRT

; ORGANISM: Candida albicans

US-09-248-796A-25823

Query Match 71.2%; Score 42; DB 4; Length 70;
Best Local Similarity 63.6%; Pred. No. 91;
Matches 7; Conservative 1; Mismatches 3; Indels 0; Gaps 0;

Qy 1 CCCSSSSSSSS 11
Db 37 CCCSSVSQNES 47

RESULT 14

US-09-270-767-39762

; Sequence 39762, Application US/09270767

; Patent No. 6703491

; GENERAL INFORMATION:

; APPLICANT: Homburger et al.

; TITLE OF INVENTION: Nucleic acids and proteins of Drosophila melanogaster

; FILE REFERENCE: File Reference: 7326-094

; CURRENT APPLICATION NUMBER: US/09/270,767

; CURRENT FILING DATE: 1999-03-17

; NUMBER OF SEQ ID NOS: 62517

; SOFTWARE: PatentIn Ver. 2.0

; SEQ ID NO 39762

; LENGTH: 148

; TYPE: PRT

; ORGANISM: Drosophila melanogaster

; FEATURE:

; OTHER INFORMATION: Xaa means any amino acid

US-09-270-767-39762

Query Match 71.2%; Score 42; DB 4; Length 148;
Best Local Similarity 87.5%; Pred. No. 1.7e+02;
Matches 7; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

Qy 1 CCCSSSSSS 8
Db 13 CCCSSRSS 20

RESULT 15

US-09-270-767-54979

; Sequence 54979, Application US/09270767

; Patent No. 6703491

; GENERAL INFORMATION:

; APPLICANT: Homburger et al.

; TITLE OF INVENTION: Nucleic acids and proteins of Drosophila melanogaster

; FILE REFERENCE: File Reference: 7326-094

; CURRENT APPLICATION NUMBER: US/09/270,767

; CURRENT FILING DATE: 1999-03-17

; NUMBER OF SEQ ID NOS: 62517

; SOFTWARE: PatentIn Ver. 2.0

; SEQ ID NO 54979

; LENGTH: 148

; TYPE: PRT

; ORGANISM: Drosophila melanogaster

; FEATURE:

; OTHER INFORMATION: Xaa means any amino acid

US-09-270-767-54979

Query Match 71.2%; Score 42; DB 4; Length 148;
Best Local Similarity 87.5%; Pred. No. 1.7e+02;
Matches 7; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

Qy 1 CCCSSSSSS 8
Db 13 CCCSSRSS 20

Search completed: November 15, 2005, 12:29:35

Job time : 22.2895 secs

This Page Blank (uspto)

GenCore version 5.1.6
Copyright (c) 1993 - 2005 Compugen Ltd.

OM protein - protein search, using sw model

Run on: November 15, 2005, 12:07:58 ; Search time 87.1316 Seconds
(without alignments)
48.827 Million cell updates/sec

Title: US-10-820-998-3

Perfect score: 59

Sequence: 1 CCCSSSSSSSS 11

Scoring table:

BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 2105692 seqs, 386760381 residues

Total number of hits satisfying chosen parameters: 2105692

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : A_Geneseq_16Dec04:*
1: Geneseqp1980s:*
2: Geneseqp1990s:*
3: Geneseqp2000s:*
4: Geneseqp2001s:*
5: Geneseqp2002s:*
6: Geneseqp2003as:*
7: Geneseqp2003bs:*
8: Geneseqp2004s:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	56	94.9	158	3	AAG29118 Arabidops
2	56	94.9	201	3	AAG29117 Arabidops
3	56	94.9	268	3	AAG29116 Arabidops
4	56	94.9	268	7	ADC46627 Thalecres
5	56	94.9	268	8	ADO01923 Thalecres
6	53	89.8	96	4	ABG27789 Novel hum
7	51	86.4	103	4	ABG28278 Novel hum
8	50	84.7	120	5	ADK34141 Novel hum
9	50	84.7	968	6	AAE38440 Human FIS
10	50	84.7	968	4	AAU14174 Human nov
11	50	84.7	1054	6	ADA55008 Human pro
12	48	81.4	91	3	AG41076 Zea mays
13	47	79.7	109	4	AAU18125 Novel hum
14	47	79.7	109	4	AAU18675 Renal and
15	47	79.7	109	4	ABU10317 Human cDN
16	47	79.7	109	4	AAU87616 Novel cen
17	47	79.7	109	4	AAU18486 Human end
18	47	79.7	109	4	ADM19684 Protein e
19	47	79.7	109	5	ABJ05752 Novel hum
20	47	79.7	109	5	ABP66904 Human pol
21	47	79.7	109	6	ABU97290 Human pol
22	47	79.7	109	8	ADI54931 Novel hum
23	47	79.7	1307	5	ABB91506 Herbicida
24	46	78.0	39	4	AAU15525 Peptide #
25	46	78.0	39	4	ABB34531 Peptide #

26	46	78.0	39	4	AAW28014	Am28014 Peptide #
27	46	78.0	39	4	ABB29358	Abb29358 Peptide #
28	46	78.0	39	4	ABBI9941	Abbi9941 Protein #
29	46	78.0	39	4	AAW67715	Am67715 Human bon
30	46	78.0	39	4	AAW55318	Am55318 Human bon
31	46	78.0	39	4	ABG49352	Abg49352 Human liv
32	46	78.0	39	4	AAW03276	Am03276 Peptide #
33	46	78.0	39	5	ABG37278	Abg37278 Human pep
34	46	78.0	53	1	AAW90361	Ap90361 Recombina
35	46	78.0	53	1	AAW90114	Ap90114 N-Termina
36	46	78.0	53	2	AAW72601	Aw72601 Group A E
37	46	78.0	99	4	ABB69461	Abb69461 Drosophil
38	46	78.0	143	7	ABO69539	AbO69539 Pseudomon
39	46	78.0	639	7	ADF43578	Adf43578 Thale cre
40	46	78.0	1987	4	ABB61083	Abb61083 Drosophil
41	45	76.3	415	4	AAU48093	Am48093 Propionib
42	45	76.3	415	6	ABM44612	Abm44612 Propionib
43	45	76.3	1077	5	ABP74094	Abp74094 Candida a
44	44	74.6	63	5	ABP33047	Abp33047 Human ORF
45	44	74.6	82	4	AAU53315	Aau53315 Propionib

ALIGNMENTS

RESULT 1
AAG29118
ID AAG29118 standard; protein; 158 AA.

XX AC AAG29118;
XX DT 17-OCT-2000 (first entry)
XX DE Arabidopsis thaliana protein fragment SEQ ID NO: 34588.
XX KW Protein identification; signal transduction pathway; metabolic pathway;
XX KW hybridisation assay; genetic mapping; gene expression control; promoter;
XX KW termination sequence.

OS Arabidopsis thaliana.

PN EP1033405-A2.

XX PD 06-SEP-2000.

XX PF 25-FEB-2000; 2000EP-00301439.

XX PR 25-FEB-1999; 99US-0121825P.

PR 05-MAR-1999; 99US-0123180P.

PR 09-MAR-1999; 99US-0123548P.

PR 23-MAR-1999; 99US-0125788P.

PR 25-MAR-1999; 99US-0126264P.

PR 29-MAR-1999; 99US-0126785P.

PR 01-APR-1999; 99US-0127462P.

PR 06-APR-1999; 99US-0128234P.

PR 08-APR-1999; 99US-0128714P.

PR 16-APR-1999; 99US-0129845P.

PR 19-APR-1999; 99US-0130077P.

PR 21-APR-1999; 99US-0130449P.

PR 23-APR-1999; 99US-0130510P.

PR 28-APR-1999; 99US-0130891P.

PR 30-APR-1999; 99US-0131449P.

PR 30-APR-1999; 99US-0132048P.

PR 30-APR-1999; 99US-0132407P.

PR 04-MAY-1999; 99US-0132484P.

PR 05-MAY-1999; 99US-0132485P.

PR 06-MAY-1999; 99US-0132486P.

PR 07-MAY-1999; 99US-0132487P.

PR 11-MAY-1999; 99US-0132863P.

PR 14-MAY-1999; 99US-0134256P.

PR 14-MAY-1999; 99US-0134218P.

PR 14-MAY-1999; 99US-0134219P.

PR 14-MAY-1999; 99US-0134221P.

PR 14-MAY-1999; 99US-0134370P. 27-JUL-1999; 99US-0145918P.
PR 18-MAY-1999; 99US-0134768P. PR 27-JUL-1999; 99US-0145919P.
PR 19-MAY-1999; 99US-0134941P. PR 28-JUL-1999; 99US-0145951P.
PR 20-MAY-1999; 99US-0135124P. PR 02-AUG-1999; 99US-0146386P.
PR 21-MAY-1999; 99US-0135353P. PR 02-AUG-1999; 99US-0146388P.
PR 24-MAY-1999; 99US-0135629P. PR 03-AUG-1999; 99US-0146389P.
PR 25-MAY-1999; 99US-0136021P. PR 03-AUG-1999; 99US-0147038P.
PR 27-MAY-1999; 99US-0136332P. PR 04-AUG-1999; 99US-0147204P.
PR 28-MAY-1999; 99US-0136782P. PR 04-AUG-1999; 99US-0147302P.
PR 01-JUN-1999; 99US-0137222P. PR 05-AUG-1999; 99US-0147192P.
PR 03-JUN-1999; 99US-0137528P. PR 05-AUG-1999; 99US-0147260P.
PR 04-JUN-1999; 99US-0137502P. PR 06-AUG-1999; 99US-0147303P.
PR 07-JUN-1999; 99US-0137724P. PR 06-AUG-1999; 99US-0147416P.
PR 08-JUN-1999; 99US-0138094P. PR 09-AUG-1999; 99US-0147493P.
PR 10-JUN-1999; 99US-0138540P. PR 09-AUG-1999; 99US-0147915P.
PR 10-JUN-1999; 99US-0138847P. PR 10-AUG-1999; 99US-0148171P.
PR 14-JUN-1999; 99US-0139119P. PR 11-AUG-1999; 99US-0148319P.
PR 16-JUN-1999; 99US-0139452P. PR 12-AUG-1999; 99US-0148341P.
PR 16-JUN-1999; 99US-0139453P. PR 13-AUG-1999; 99US-0148565P.
PR 17-JUN-1999; 99US-0139492P. PR 13-AUG-1999; 99US-0148684P.
PR 18-JUN-1999; 99US-0139454P. PR 16-AUG-1999; 99US-0149168P.
PR 18-JUN-1999; 99US-0139455P. PR 17-AUG-1999; 99US-0149175P.
PR 18-JUN-1999; 99US-0139456P. PR 18-AUG-1999; 99US-0149426P.
PR 18-JUN-1999; 99US-0139457P. PR 20-AUG-1999; 99US-0149722P.
PR 18-JUN-1999; 99US-0139458P. PR 20-AUG-1999; 99US-0149723P.
PR 18-JUN-1999; 99US-0139459P. PR 23-AUG-1999; 99US-0149929P.
PR 18-JUN-1999; 99US-0139460P. PR 23-AUG-1999; 99US-0149929P.
PR 18-JUN-1999; 99US-0139461P. PR 25-AUG-1999; 99US-0149929P.
PR 18-JUN-1999; 99US-0139462P. PR 26-AUG-1999; 99US-0150566P.
PR 18-JUN-1999; 99US-0139463P. PR 27-AUG-1999; 99US-0150884P.
PR 18-JUN-1999; 99US-0139750P. PR 27-AUG-1999; 99US-0151065P.
PR 18-JUN-1999; 99US-0139763P. PR 27-AUG-1999; 99US-0151080P.
PR 21-JUN-1999; 99US-0139817P. PR 30-AUG-1999; 99US-0151303P.
PR 22-JUN-1999; 99US-0139899P. PR 31-AUG-1999; 99US-0151438P.
PR 23-JUN-1999; 99US-0140353P. PR 01-SEP-1999; 99US-0151930P.
PR 23-JUN-1999; 99US-0140354P. PR 07-SEP-1999; 99US-0152363P.
PR 24-JUN-1999; 99US-0140695P. PR 10-SEP-1999; 99US-0153070P.
PR 28-JUN-1999; 99US-0140821P. PR 13-SEP-1999; 99US-0153758P.
PR 29-JUN-1999; 99US-0140991P. PR 15-SEP-1999; 99US-0154018P.
PR 30-JUN-1999; 99US-0141287P. PR 16-SEP-1999; 99US-0154039P.
PR 01-JUL-1999; 99US-0141842P. PR 20-SEP-1999; 99US-0154779P.
PR 01-JUL-1999; 99US-0142154P. PR 22-SEP-1999; 99US-0155139P.
PR 02-JUL-1999; 99US-0142055P. PR 23-SEP-1999; 99US-0155486P.
PR 06-JUL-1999; 99US-0142330P. PR 24-SEP-1999; 99US-0155659P.
PR 08-JUL-1999; 99US-0142803P. PR 28-SEP-1999; 99US-0156458P.
PR 09-JUL-1999; 99US-0142920P. PR 29-SEP-1999; 99US-0156596P.
PR 12-JUL-1999; 99US-0142977P. PR 04-OCT-1999; 99US-0157117P.
PR 13-JUL-1999; 99US-0143542P. PR 05-OCT-1999; 99US-0157753P.
PR 14-JUL-1999; 99US-0143624P. PR 06-OCT-1999; 99US-0157865P.
PR 15-JUL-1999; 99US-0144005P. PR 07-OCT-1999; 99US-0158029P.
PR 16-JUL-1999; 99US-0144085P. PR 08-OCT-1999; 99US-0158222P.
PR 16-JUL-1999; 99US-0144086P. PR 12-OCT-1999; 99US-0158369P.
PR 19-JUL-1999; 99US-0144325P. PR 13-OCT-1999; 99US-0159293P.
PR 19-JUL-1999; 99US-0144331P. PR 13-OCT-1999; 99US-0159294P.
PR 19-JUL-1999; 99US-0144332P. PR 13-OCT-1999; 99US-0159295P.
PR 19-JUL-1999; 99US-0144333P. PR 14-OCT-1999; 99US-0159329P.
PR 19-JUL-1999; 99US-0144334P. PR 14-OCT-1999; 99US-0159330P.
PR 19-JUL-1999; 99US-0144335P. PR 14-OCT-1999; 99US-0159331P.
PR 20-JUL-1999; 99US-0144352P. PR 14-OCT-1999; 99US-0159637P.
PR 20-JUL-1999; 99US-0144632P. PR 14-OCT-1999; 99US-0159638P.
PR 20-JUL-1999; 99US-0144884P. PR 18-OCT-1999; 99US-0159584P.
PR 21-JUL-1999; 99US-0144814P. PR 21-OCT-1999; 99US-0160741P.
PR 21-JUL-1999; 99US-0145086P. PR 21-OCT-1999; 99US-0160767P.
PR 21-JUL-1999; 99US-0145088P. PR 21-OCT-1999; 99US-0160768P.
PR 22-JUL-1999; 99US-0145085P. PR 21-OCT-1999; 99US-0160770P.
PR 22-JUL-1999; 99US-0145087P. PR 21-OCT-1999; 99US-0160814P.
PR 22-JUL-1999; 99US-0145089P. PR 21-OCT-1999; 99US-0160815P.
PR 22-JUL-1999; 99US-0145192P. PR 22-OCT-1999; 99US-0160980P.
PR 23-JUL-1999; 99US-0145145P. PR 22-OCT-1999; 99US-0160981P.
PR 23-JUL-1999; 99US-0145218P. PR 22-OCT-1999; 99US-0160989P.
PR 23-JUL-1999; 99US-0145224P. PR 25-OCT-1999; 99US-0161404P.
PR 26-JUL-1999; 99US-0145276P. PR 25-OCT-1999; 99US-0161405P.
PR 27-JUL-1999; 99US-0145913P.

```
PR 25-OCT-1999; 99US-0161406P.
PR 26-OCT-1999; 99US-0161359P.
PR 26-OCT-1999; 99US-0161360P.
PR 26-OCT-1999; 99US-0161361P.
PR 28-OCT-1999; 99US-0161920P.
PR 28-OCT-1999; 99US-0161922P.
PR 28-OCT-1999; 99US-0161933P.
PR 29-OCT-1999; 99US-0162142P.

Query Match          94.9%; Score 56; DB 3; Length 158;
Best Local Similarity 90.9%; Pred. No. 13;
Matches 10; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

Qy 1 CCCSSSSSSSS 11
Db 119 CCCSSSTSSSS 129

RESULT 2
AAG29117
ID AAG29117 standard; protein; 201 AA.
XX
AC AAG29117;
XX
DT 17-OCT-2000 (first entry)
XX
DE Arabidopsis thaliana protein fragment SEQ ID NO: 34587.
XX
KW Protein identification; signal transduction pathway; metabolic pathway;
KW hybridisation assay; genetic mapping; gene expression control; promoter;
KW termination sequence.
XX
OS Arabidopsis thaliana.
XX
PN EP1033405-A2.
XX
PD 06-SEP-2000.
XX
PF 25-FEB-2000; 2000EP-00301439.
XX
PR 25-FEB-1999; 99US-0121825P.
PR 05-MAR-1999; 99US-0123180P.
PR 09-MAR-1999; 99US-0123548P.
PR 23-MAR-1999; 99US-0125788P.
PR 25-MAR-1999; 99US-0126264P.
PR 29-MAR-1999; 99US-0126785P.
PR 01-APR-1999; 99US-0127462P.
PR 06-APR-1999; 99US-0128234P.
PR 08-APR-1999; 99US-0128714P.
PR 16-APR-1999; 99US-0129845P.
PR 19-APR-1999; 99US-0130077P.
PR 21-APR-1999; 99US-0130449P.
PR 23-APR-1999; 99US-0130510P.
PR 28-APR-1999; 99US-0131449P.
PR 30-APR-1999; 99US-0132048P.
PR 04-MAY-1999; 99US-0132407P.
PR 05-MAY-1999; 99US-0132484P.
PR 06-MAY-1999; 99US-0132485P.
PR 06-MAY-1999; 99US-0132486P.
PR 07-MAY-1999; 99US-0132487P.
PR 11-MAY-1999; 99US-0132863P.
PR 14-MAY-1999; 99US-0134218P.
PR 14-MAY-1999; 99US-0134219P.
PR 14-MAY-1999; 99US-0134221P.
PR 14-MAY-1999; 99US-0134370P.
PR 18-MAY-1999; 99US-0134768P.
PR 19-MAY-1999; 99US-0134941P.
PR 20-MAY-1999; 99US-0135124P.
PR 21-MAY-1999; 99US-0135353P.
PR 24-MAY-1999; 99US-0135629P.
PR 25-MAY-1999; 99US-0136021P.

99US-0136392P.
99US-0136782P.
99US-0137222P.
99US-0137528P.
99US-0137502P.
99US-0137724P.
99US-0138094P.
99US-0138540P.
99US-0138847P.
99US-0139119P.
99US-0139452P.
99US-0139453P.
99US-0139492P.
99US-0139454P.
99US-0139455P.
99US-0139456P.
99US-0139457P.
99US-0139458P.
99US-0139460P.
99US-0139461P.
99US-0139462P.
99US-0139463P.
99US-0139750P.
99US-0139763P.
99US-0139817P.
99US-0139899P.
99US-0140353P.
99US-0140354P.
99US-0140695P.
99US-0140823P.
99US-0140991P.
99US-0141287P.
99US-0141842P.
99US-0142154P.
99US-0142055P.
99US-0142390P.
99US-0142803P.
99US-0142920P.
99US-0142977P.
99US-0143542P.
99US-0143624P.
99US-0144005P.
99US-0144085P.
99US-0144086P.
99US-0144325P.
99US-0144331P.
99US-0144332P.
99US-0144333P.
99US-0144334P.
99US-0144335P.
99US-0144352P.
99US-0144632P.
99US-0144884P.
99US-0144814P.
99US-0145086P.
99US-0145088P.
99US-0145089P.
99US-0145087P.
99US-0145089P.
99US-0145192P.
99US-0145145P.
99US-0145218P.
99US-0145224P.
99US-0145276P.
99US-0145913P.
99US-0145918P.
99US-0145919P.
99US-0145951P.
99US-0146386P.
99US-0146388P.
99US-0146389P.
99US-0147038P.
```

```
PR 04-AUG-1999; 99US-0147204P.
PR 04-AUG-1999; 99US-0147302P.
PR 05-AUG-1999; 99US-0147192P.
PR 05-AUG-1999; 99US-0147260P.
PR 06-AUG-1999; 99US-0147303P.
PR 06-AUG-1999; 99US-0147416P.
PR 09-AUG-1999; 99US-0147493P.
PR 09-AUG-1999; 99US-0147935P.
PR 10-AUG-1999; 99US-0148171P.
PR 11-AUG-1999; 99US-0148319P.
PR 12-AUG-1999; 99US-0148341P.
PR 13-AUG-1999; 99US-0148565P.
PR 13-AUG-1999; 99US-0148684P.
PR 16-AUG-1999; 99US-0149368P.
PR 17-AUG-1999; 99US-0149175P.
PR 18-AUG-1999; 99US-0149426P.
PR 20-AUG-1999; 99US-0149722P.
PR 20-AUG-1999; 99US-0149723P.
PR 20-AUG-1999; 99US-0149929P.
PR 23-AUG-1999; 99US-0149902P.
PR 23-AUG-1999; 99US-0149930P.
PR 25-AUG-1999; 99US-0150566P.
PR 26-AUG-1999; 99US-0150884P.
PR 27-AUG-1999; 99US-0151065P.
PR 27-AUG-1999; 99US-0151066P.
PR 27-AUG-1999; 99US-0151080P.
PR 30-AUG-1999; 99US-0151303P.
PR 31-AUG-1999; 99US-0151438P.
PR 01-SEP-1999; 99US-0151930P.
PR 07-SEP-1999; 99US-0152363P.
PR 10-SEP-1999; 99US-0153070P.
PR 13-SEP-1999; 99US-0153758P.
PR 15-SEP-1999; 99US-0154018P.
PR 16-SEP-1999; 99US-0154039P.
PR 20-SEP-1999; 99US-0154779P.
PR 22-SEP-1999; 99US-0155139P.
PR 23-SEP-1999; 99US-0155486P.
PR 24-SEP-1999; 99US-0155659P.
PR 28-SEP-1999; 99US-0156458P.
PR 29-SEP-1999; 99US-0156596P.
PR 04-OCT-1999; 99US-0157117P.
PR 05-OCT-1999; 99US-0157753P.
PR 06-OCT-1999; 99US-0157865P.
PR 07-OCT-1999; 99US-0158029P.
PR 08-OCT-1999; 99US-0158232P.
PR 12-OCT-1999; 99US-0158369P.
PR 13-OCT-1999; 99US-0159293P.
PR 13-OCT-1999; 99US-0159294P.
PR 13-OCT-1999; 99US-0159295P.
PR 14-OCT-1999; 99US-0159339P.
PR 14-OCT-1999; 99US-0159330P.
PR 14-OCT-1999; 99US-0159331P.
PR 14-OCT-1999; 99US-0159637P.
PR 14-OCT-1999; 99US-0159638P.
PR 18-OCT-1999; 99US-0159594P.
PR 21-OCT-1999; 99US-0160741P.
PR 21-OCT-1999; 99US-0160767P.
PR 21-OCT-1999; 99US-0160768P.
PR 21-OCT-1999; 99US-0160770P.
PR 21-OCT-1999; 99US-0160814P.
PR 21-OCT-1999; 99US-0160815P.
PR 22-OCT-1999; 99US-0160980P.
PR 22-OCT-1999; 99US-0160981P.
PR 22-OCT-1999; 99US-0160989P.
PR 25-OCT-1999; 99US-0161404P.
PR 25-OCT-1999; 99US-0161405P.
PR 25-OCT-1999; 99US-0161406P.
PR 26-OCT-1999; 99US-0161359P.
PR 26-OCT-1999; 99US-0161360P.
PR 26-OCT-1999; 99US-0161361P.
PR 28-OCT-1999; 99US-0161920P.
PR 28-OCT-1999; 99US-0161922P.
PR 28-OCT-1999; 99US-0161933P.

PR 29-OCT-1999; 99US-0162142P.
Query Match 94.9%; Score 56; DB 3; Length 201;
Best Local Similarity 90.9%; Pred. No. 16;
Matches 10; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

QY 1 CCCSSSSSSSS 11
Db 162 CCCSSSSSSSS 172

RESULT 3
AAG29116
ID AAG29116 standard; protein; 268 AA.
XX
AC AAG29116;
XX
DT 17-OCT-2000 (first entry)
XX
DE Arabidopsis thaliana protein fragment SEQ ID NO: 34586.
XX
KW Protein identification; signal transduction pathway; metabolic pathway;
hybridisation assay; genetic mapping; gene expression control; promoter;
termination sequence.
XX
OS Arabidopsis thaliana.
XX
PN EP1033405-A2.
PD 06-SEP-2000.
PF
XX
XX 25-FEB-2000; 2000EP-00301439.
PR 25-FEB-1999; 99US-0121825P.
PR 05-MAR-1999; 99US-0123180P.
PR 09-MAR-1999; 99US-0123548P.
PR 23-MAR-1999; 99US-0125788P.
PR 25-MAR-1999; 99US-0126264P.
PR 29-MAR-1999; 99US-0126785P.
PR 01-APR-1999; 99US-0127462P.
PR 06-APR-1999; 99US-0128234P.
PR 08-APR-1999; 99US-0128714P.
PR 16-APR-1999; 99US-0129845P.
PR 21-APR-1999; 99US-0130077P.
PR 23-APR-1999; 99US-0130449P.
PR 23-APR-1999; 99US-0130510P.
PR 28-APR-1999; 99US-0130891P.
PR 28-APR-1999; 99US-0131449P.
PR 30-APR-1999; 99US-0132048P.
PR 30-APR-1999; 99US-0132407P.
PR 04-MAY-1999; 99US-0132484P.
PR 05-MAY-1999; 99US-0132485P.
PR 06-MAY-1999; 99US-0132486P.
PR 06-MAY-1999; 99US-0132487P.
PR 07-MAY-1999; 99US-0132863P.
PR 11-MAY-1999; 99US-0134256P.
PR 14-MAY-1999; 99US-0134218P.
PR 14-MAY-1999; 99US-0134219P.
PR 14-MAY-1999; 99US-0134421P.
PR 14-MAY-1999; 99US-0134370P.
PR 18-MAY-1999; 99US-0134768P.
PR 19-MAY-1999; 99US-0134941P.
PR 20-MAY-1999; 99US-0135124P.
PR 21-MAY-1999; 99US-0135353P.
PR 21-MAY-1999; 99US-0135629P.
PR 25-MAY-1999; 99US-0136021P.
PR 27-MAY-1999; 99US-0136392P.
PR 28-MAY-1999; 99US-0136782P.
PR 01-JUN-1999; 99US-0137222P.
PR 03-JUN-1999; 99US-0137528P.
PR 04-JUN-1999; 99US-0137502P.
PR 07-JUN-1999; 99US-0137724P.
PR 08-JUN-1999; 99US-0138094P.
```


PR 10-JUN-1999; 99US-0138540P.
PR 10-JUN-1999; 99US-0138847P.
PR 14-JUN-1999; 99US-0139119P.
PR 16-JUN-1999; 99US-0139452P.
PR 16-JUN-1999; 99US-0139453P.
PR 17-JUN-1999; 99US-0139452P.
PR 18-JUN-1999; 99US-0139454P.
PR 18-JUN-1999; 99US-0139455P.
PR 18-JUN-1999; 99US-0139456P.
PR 18-JUN-1999; 99US-0139457P.
PR 18-JUN-1999; 99US-0139458P.
PR 18-JUN-1999; 99US-0139459P.
PR 18-JUN-1999; 99US-0139460P.
PR 18-JUN-1999; 99US-0139461P.
PR 18-JUN-1999; 99US-0139462P.
PR 18-JUN-1999; 99US-0139463P.
PR 18-JUN-1999; 99US-0139750P.
PR 18-JUN-1999; 99US-0139763P.
PR 21-JUN-1999; 99US-0139817P.
PR 22-JUN-1999; 99US-0139899P.
PR 23-JUN-1999; 99US-0140353P.
PR 23-JUN-1999; 99US-0140354P.
PR 24-JUN-1999; 99US-0140695P.
PR 28-JUN-1999; 99US-0140823P.
PR 29-JUN-1999; 99US-0140991P.
PR 30-JUN-1999; 99US-0141287P.
PR 01-JUL-1999; 99US-0141842P.
PR 01-JUL-1999; 99US-0142154P.
PR 02-JUL-1999; 99US-0142055P.
PR 08-JUL-1999; 99US-0142390P.
PR 08-JUL-1999; 99US-0142803P.
PR 09-JUL-1999; 99US-0142920P.
PR 12-JUL-1999; 99US-0142977P.
PR 13-JUL-1999; 99US-0143542P.
PR 14-JUL-1999; 99US-0143624P.
PR 15-JUL-1999; 99US-0144005P.
PR 16-JUL-1999; 99US-0144085P.
PR 16-JUL-1999; 99US-0144086P.
PR 19-JUL-1999; 99US-0144325P.
PR 19-JUL-1999; 99US-0144331P.
PR 19-JUL-1999; 99US-0144332P.
PR 19-JUL-1999; 99US-0144333P.
PR 19-JUL-1999; 99US-0144334P.
PR 19-JUL-1999; 99US-0144335P.
PR 20-JUL-1999; 99US-0144352P.
PR 20-JUL-1999; 99US-0144632P.
PR 20-JUL-1999; 99US-0144884P.
PR 21-JUL-1999; 99US-0144814P.
PR 21-JUL-1999; 99US-0145086P.
PR 21-JUL-1999; 99US-0145088P.
PR 22-JUL-1999; 99US-0145087P.
PR 22-JUL-1999; 99US-0145089P.
PR 22-JUL-1999; 99US-0145192P.
PR 23-JUL-1999; 99US-0145145P.
PR 23-JUL-1999; 99US-0145218P.
PR 23-JUL-1999; 99US-0145224P.
PR 26-JUL-1999; 99US-0145276P.
PR 27-JUL-1999; 99US-0145913P.
PR 27-JUL-1999; 99US-0145918P.
PR 27-JUL-1999; 99US-0145919P.
PR 28-JUL-1999; 99US-0145951P.
PR 28-JUL-1999; 99US-0146386P.
PR 02-AUG-1999; 99US-0146388P.
PR 02-AUG-1999; 99US-0146389P.
PR 03-AUG-1999; 99US-0147038P.
PR 04-AUG-1999; 99US-0147204P.
PR 04-AUG-1999; 99US-0147302P.
PR 05-AUG-1999; 99US-0147192P.
PR 05-AUG-1999; 99US-0147260P.
PR 06-AUG-1999; 99US-0147303P.
PR 06-AUG-1999; 99US-0147416P.
PR 09-AUG-1999; 99US-0147493P.

PR 09-AUG-1999; 99US-0147935P.
PR 10-AUG-1999; 99US-0148171P.
PR 11-AUG-1999; 99US-0148319P.
PR 12-AUG-1999; 99US-0148341P.
PR 13-AUG-1999; 99US-0148565P.
PR 13-AUG-1999; 99US-0148684P.
PR 16-AUG-1999; 99US-0149368P.
PR 17-AUG-1999; 99US-0149175P.
PR 18-AUG-1999; 99US-0149426P.
PR 20-AUG-1999; 99US-0149722P.
PR 20-AUG-1999; 99US-0149723P.
PR 20-AUG-1999; 99US-0149929P.
PR 23-AUG-1999; 99US-0149902P.
PR 23-AUG-1999; 99US-0149930P.
PR 25-AUG-1999; 99US-0150566P.
PR 26-AUG-1999; 99US-0150884P.
PR 27-AUG-1999; 99US-0151065P.
PR 27-AUG-1999; 99US-0151066P.
PR 27-AUG-1999; 99US-0151080P.
PR 30-AUG-1999; 99US-0151303P.
PR 31-AUG-1999; 99US-0151438P.
PR 01-SEP-1999; 99US-0151930P.
PR 07-SEP-1999; 99US-0152363P.
PR 10-SEP-1999; 99US-0153070P.
PR 13-SEP-1999; 99US-0153758P.
PR 15-SEP-1999; 99US-0154018P.
PR 16-SEP-1999; 99US-0154039P.
PR 20-SEP-1999; 99US-0154779P.
PR 22-SEP-1999; 99US-0155139P.
PR 23-SEP-1999; 99US-0155486P.
PR 24-SEP-1999; 99US-0155659P.
PR 28-SEP-1999; 99US-0156458P.
PR 29-SEP-1999; 99US-0156596P.
PR 04-OCT-1999; 99US-0157117P.
PR 05-OCT-1999; 99US-0157753P.
PR 06-OCT-1999; 99US-0157865P.
PR 07-OCT-1999; 99US-0158029P.
PR 08-OCT-1999; 99US-0158232P.
PR 12-OCT-1999; 99US-0158369P.
PR 13-OCT-1999; 99US-0159293P.
PR 13-OCT-1999; 99US-0159294P.
PR 13-OCT-1999; 99US-0159295P.
PR 14-OCT-1999; 99US-0159329P.
PR 14-OCT-1999; 99US-0159330P.
PR 14-OCT-1999; 99US-0159331P.
PR 14-OCT-1999; 99US-0159637P.
PR 14-OCT-1999; 99US-0159638P.
PR 18-OCT-1999; 99US-0159584P.
PR 21-OCT-1999; 99US-0160741P.
PR 21-OCT-1999; 99US-0160768P.
PR 21-OCT-1999; 99US-0160770P.
PR 21-OCT-1999; 99US-0160814P.
PR 21-OCT-1999; 99US-0160815P.
PR 22-OCT-1999; 99US-0160980P.
PR 22-OCT-1999; 99US-0160981P.
PR 22-OCT-1999; 99US-0160989P.
PR 25-OCT-1999; 99US-0161404P.
PR 25-OCT-1999; 99US-0161405P.
PR 25-OCT-1999; 99US-0161406P.
PR 26-OCT-1999; 99US-0161359P.
PR 26-OCT-1999; 99US-0161360P.
PR 26-OCT-1999; 99US-0161361P.
PR 28-OCT-1999; 99US-0161920P.
PR 28-OCT-1999; 99US-0161922P.
PR 28-OCT-1999; 99US-0161993P.
PR 29-OCT-1999; 99US-0162142P.

Query Match 94.9%; Score 56; DB 3; Length 268;
Best Local Similarity 90.9%; Pred. No. 20;
Matches 10; Conservative 1; Mismatches 0; Indels 0; Gaps 0;
QY 1 CCCSSSSSSSS 11

```

Db          229 CCCSSTSSSS 239
          |||||:||||
RESULT 4
ADC46627
ID   ADC46627 standard; protein; 268 AA.
XX   AC
XX   AC ADC46627;
XX   DT 18-DEC-2003 (first entry)
XX   DE Thalecress transcription factor-like protein G524.
XX   KW Thalecress; transcription factor-like protein; seed trait; transgenic;
XX   KW plant size; stress tolerance; yield; disease resistance; plant.
XX   OS Arabidopsis thaliana.
XX   PN US2003093837-A1.
XX   PD 15-MAY-2003.
XX   PF 01-NOV-2002; 2002US-00286264.
XX   PR 23-MAR-1999; 99US-0125814P.
XX   PR 22-MAR-2000; 2000US-00533030.
XX   PA (KEDD/) KEDDIE J.
XX   PA (RIEC/) RIECHMANN J L.
XX   PA (RATC/) RATCLIFFE O.
XX   PA (ZHAN/) ZHANG J.
XX   PA (JIAN/) JIANG C.
XX   PA (PINE/) PINEDA O.
XX   PA (HEAR/) HEARD J.
XX   PA (YUGG/) YU G.
XX   PA (ADAM/) ADAM L.
XX   PA (BROU/) BROUN P.
XX   PA (REUB/) REUBER L.
XX   PA (PLIG/) PILGRIM M.
XX   PA (SAMA/) SAMAHA R.
XX   PI Keddie J, Riechmann JH, Ratcliffe O, Zhang J, Jiang C, Pineda O;
XX   PI Heard J, Yu G, Adam L, Broun P, Reuber L, Pilgrim M, Samaha R;
XX   DR MPI; 2003-765498/72.
XX   DR N-PSDB; ADC46626.
XX   PT Novel transgenic plant having recombinant polynucleotide encoding
XX   PT polypeptide that alters trait of transgenic plant when compared with same
XX   PT trait of another plant lacking recombinant polynucleotide.
XX   PS Claim 1; SEQ ID NO 26; 165pp; English.
XX   CC The invention relates to a transgenic plant having recombinant
XX   CC polynucleotide (II) encoding polypeptide comprising at least 6
XX   CC consecutive amino acids of a sequence chosen from the protein sequence
XX   CC appearing as ADC46603 - ADC46749 (every second sequence), where
XX   CC recombinant polynucleotide alters a trait of the seed transgenic plant
XX   CC when compared with same trait of another plant lacking recombinant
XX   CC polynucleotide. The proteins are transcription factor-like proteins. Also
XX   CC included are altering (M1) a trait associated with seed (comprising:
XX   CC transforming a plant with (II); selecting the transformed plants; and
XX   CC identifying a transformed plant with seed having altered trait), altering
XX   CC (M2) the expression levels of at least one gene of a plant (involving
XX   CC transforming the plant with (II) and selecting the transformed plant),
XX   CC altering (M3) a trait associated with a plant's seed (comprising:
XX   CC transforming the plant with a recombinant polynucleotide comprising a
XX   CC nucleotide sequence comprising at least 18 consecutive nucleotides of a
XX   CC sequence appearing as ADC46750 - ADC46766 and selecting the transformed
XX   CC plan) altering (M4) a plant's trait (involving providing a database
XX   CC sequence, comparing the database sequence with a polypeptide or a
XX   CC polynucleotide chosen as detailed above, selecting a database sequence

```

```

CC that needs selected sequence criteria and transforming a database
CC sequence in the plant) and altering a plant's trait (involving providing
CC a test polynucleotide, hybridising the test polynucleotide with a
CC polynucleotide detailed above at low stringency and transforming the
CC hybridising test polynucleotide in a plant to alter a trait of the
CC plant). The method (M1) is useful for altering a trait associated with
CC seed. The method (M2) is useful for altering the expressing levels of at
CC least one gene of a plant. The method (M3) is useful for altering a trait
CC associated with a plant's seed. The method (M4) is useful for altering a
CC plant's trait. The method (M4) is useful for altering a plant's trait
CC such as seed or plant size, stress tolerance, yield or disease
CC resistance. The present sequence represents a transcription factor-like
CC protein/seed trait altering protein of the invention.
XX
XX SQ Sequence 268 AA;
          Query Match          94.9%; Score 56; DB 7; Length 268;
          Best Local Similarity 90.9%; Pred. No. 20;
          Matches 10; Conservative 1; Mismatches 0; Indels 0; Gaps 0;
QY 1 CCCSSSSSSSS 11
Db 229 CCCSSTSSSS 239
          |||||:||||
RESULT 5
AD001923
ID   AD001923 standard; protein; 268 AA.
XX   AC
XX   AC AD001923;
XX   DT 01-JUL-2004 (first entry)
XX   DE Thalecress transcription factor protein #168.
XX   KW Thalecress; transcription factor; plant; transgenic; abiotic stress;
XX   KW cold tolerance; heat tolerance; drought; osmotic stress;
XX   KW phosphate limitation; potassium limitation; nitrogen limitation;
XX   KW hormone sensitivity; disease resistance; sugar sensing; seed germination;
XX   KW flowering; inflorescence architectural change;
XX   KW meristem cell differentiation; phyllotaxy; apical dominance;
XX   KW trichome development; seed development; premature senescence;
XX   KW delayed senescence; lethality; necrosis; plant size; leaf morphology;
XX   KW seed morphology; secondary metabolism; light response; shade avoidance.
XX   OS Arabidopsis thaliana.
XX   PN US2004045049-A1.
XX   PD 04-MAR-2004.
XX   PF 10-APR-2003; 2003US-00412699.
XX   PR 13-SEP-1999; 99US-00394519.
XX   PR 21-JAN-2000; 2000US-00489376.
XX   PR 17-FEB-2000; 2000US-00506720.
XX   PR 22-MAR-2000; 2000US-00532591.
XX   PR 22-MAR-2000; 2000US-00533029.
XX   PR 22-MAR-2000; 2000US-00533030.
XX   PR 22-MAR-2000; 2000US-00533392.
XX   PR 22-MAR-2000; 2000US-00533648.
XX   PR 06-APR-2000; 2000MO-US009448.
XX   PR 16-NOV-2000; 2000US-00713994.
XX   PR 27-MAR-2001; 2001US-00819142.
XX   PR 17-APR-2001; 2001US-00837444.
XX   PR 30-JAN-2002; 2002US-00958131.
XX   PR 14-JUN-2002; 2002US-00171468.
XX   PR 09-AUG-2002; 2002US-00225066.
XX   PR 09-AUG-2002; 2002US-00225067.
XX   PR 09-AUG-2002; 2002US-00225068.
XX   PR 17-DEC-2002; 2002US-0434166P.
XX   PR 25-FEB-2003; 2003US-00374780.
XX

```

PA (ZHAN//) ZHANG J.
PA (FROM//) FROMM M E.
PA (HEAR//) HEARD J E.
PA (RIEC//) RIECHMANN J L.
PA (ADAM//) ADAM L J.
PA (BROU//) BROUN P E.
PA (PINE//) PINEDA O.
PA (REUB//) REUBER T L.
PA (KEDD//) KEDDIE J S.
PA (YUGG//) YU G.
PA (JIAN//) JIANG C.
PA (SAMA//) SAMAHA R S.
PA (PILG//) PILGRIM M L.
PA (CREE//) CREELMAN R A.
PA (DUBE//) DUBELL A N.
PA (RATC//) RATCLIFFE O.
PA (KUMI//) KUMIMOTO R.
PA (SHER//) SHERMAN B K.
XX
PI Zhang J, Fromm ME, Heard JE, Riechmann JL, Adam LJ, Broun PE;
PI Pineda O, Reuber TL, Keddie JS, Yu G, Jiang C, Samaha RS;
PI Pilgrim ML, Creelman RA, Dubell AN, Ratcliffe O, Kumimoto R;
PI Sherman BK;
XX
DR WPI; 2004-225755/21.
DR N-PSDB; ADO01922.
XX
PT New transgenic plant, useful in developing phenotypes with altered or
PT improved characteristics or traits.
XX
PS Claim 1; SEQ ID NO 336; 213pp; English.
XX
CC The invention relates to a transgenic plant comprising a recombinant
CC polynucleotide having a polynucleotide sequence or its complementary
CC sequence comprising a sequence encoding a polypeptide, that initiates
CC transcription (i.e. a transcription factor) from Arabidopsis, Soybean,
CC Rice, Rape or corn, comprising any of the sequences appearing as ADO01588
CC -ADO03527 or ADO03530-ADO03559. Also included are using a transgenic
CC plant to grow a progeny plant, an expression cassette (comprising a
CC constitutive, inducible or tissue-specific promoter and a recombinant
CC polynucleotide described above), a host cell comprising the expression
CC cassette, producing a modified plant having a modified trait, identifying
CC a factor that is modulated by or interacts with a polypeptide encoded by
CC the polynucleotide sequence and identifying at least one downstream
CC polynucleotide sequence that is subject to a regulatory effect of any of
CC the polypeptides encoded by the polynucleotide described above. The
CC transgenic plant is useful for producing a plant that has an altered
CC trait, e.g. an enhanced tolerance to abiotic stress (increased tolerance
CC to chilling, germination in cold conditions, freezing tolerance, tolerance
CC to heat, tolerance to drought, tolerance to osmotic stress, tolerance to
CC salt, tolerance to phosphate limitation, tolerance to potassium
CC limitation, decreased sensitivity to nitrogen limitation), altered
CC hormone sensitivity, reduced sensitivity to abscisic acid, an altered
CC response to ethylene, disease resistance, altered susceptibility to
CC Botrytis, altered susceptibility to Fusarium, altered susceptibility to
CC Erysiphe, altered susceptibility to Pseudomonas syringae, altered
CC susceptibility to Sclerotinia, altered sugar sensing, improved seed
CC germination and seedling vigor, early flowering, late flowering, extended
CC period of flowering, an inflorescence architectural change, a change in
CC stem bifurcations, a lack of a shoot meristem, reduced meristem cell
CC differentiation, altered phylloclaxy, altered branching pattern, reduced
CC apical dominance, reduced trichome density, ectopic trichome development,
CC altered trichome development, altered stem morphology, increased root
CC growth, increased root hairs, altered seed development, altered cell
CC proliferation/cell differentiation, premature senescence, delayed
CC senescence, lethality, increased necrosis, an increase in seedling or
CC plant size, decreased plant size, a change in leaf morphology, increased
CC altered leaf development, increased leaf size and mass, glossy leaves,
CC leaf cell expansion, change in seed morphology, altered seed coloration,
CC increased seed size, decreased seed size, altered seed shape, change in
CC leaf biochemistry, increased leaf wax, an alteration in leaf prenyl lipid
CC content, increased leaf insoluble sugars, decreased leaf insoluble
CC sugars, increased leaf anthocyanins, an alteration of leaf fatty acid

CC content, an alteration of leaf glucosinolate content, change in seed
CC biochemistry, an increase in seed oil content, decrease in seed oil
CC content, increase in seed fatty acid content, decrease in seed fatty acid
CC content, increase in seed protein content, decrease in seed protein
CC content, alteration in seed prenyl lipid content, increase in seed
CC sterols, upregulation of genes involved in secondary metabolism, increase
CC in root anthocyanins, increase in plant anthocyanins, and alteration in
CC light response or shade avoidance. The present sequence represents a
CC thalacress transcription factor of the invention.
XX
SQ Sequence 268 AA;

Query Match 94.9%; Score 56; DB 8; Length 268;
Best Local Similarity 90.9%; Pred. No. 20;
Matches 10; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

QY 1 CCCSSSSSSSS 11
|||:|:|:|:
Db 229 CCCSSSTSSSS 239

RESULT 6
ABG27789
ID ABG27789 standard; protein; 96 AA.
XX
AC ABG27789;
XX
DT 18-FEB-2002 (first entry)
XX
DE Novel human diagnostic protein #27780.
XX
KW Human; chromosome mapping; gene mapping; gene therapy; forensic;
KW food supplement; medical imaging; diagnostic; genetic disorder.
XX
OS Homo sapiens.
XX
FN WO200175067-A2.
XX
PD 11-OCT-2001.
XX
PF 30-MAR-2001; 2001WO-US008631.
XX
PR 31-MAR-2000; 2000US-00540217.
PR 23-AUG-2000; 2000US-00649167.
XX
PA (HYSE-) HYSEQ INC.
XX
PI Drmanac RT, Liu C, Tang YT;
XX
XX WPI; 2001-639362/73.
DR N-PSDB; AAS91976.
XX
PT New isolated polynucleotide and encoded polypeptides, useful in
PT diagnostics, forensic, gene mapping, identification of mutations
PT responsible for genetic disorders or other traits and to assess
PT biodiversity.
XX
PS Claim 20; SEQ ID NO 58148; 103pp; English.
XX
CC The invention relates to isolated polynucleotide (I) and polypeptide (II)
CC sequences. (I) is useful as hybridisation probes, polymerase chain
CC reaction (PCR) primers, oligomers, and for chromosome and gene mapping,
CC and in recombinant production of (II). The polynucleotides are also used
CC in diagnostics as expressed sequence tags for identifying expressed
CC genes. (I) is useful in gene therapy techniques to restore normal
CC activity of (II) or to treat disease states involving (II). (II) is
CC useful for generating antibodies against it, detecting or quantitating a
CC polypeptide in tissue, as molecular weight markers and as a food
CC supplement. (II) and its binding partners are useful in medical imaging
CC of sites expressing (II). (I) and (II) are useful for treating disorders
CC involving aberrant protein expression or biological activity. The
CC polypeptide and polynucleotide sequences have application in
CC diagnostics, forensics, gene mapping, identification of mutations

CC responsible for genetic disorders or other traits to assess biodiversity
 CC and to produce other types of data and products dependent on DNA and
 CC amino acid sequences. ABG00010-ABG30377 represent novel human diagnostic
 CC amino acid sequences of the invention. Note: The sequence data for this
 CC patent did not appear in the printed specification, but was obtained in
 CC electronic format directly from WIPO at
 CC ftp.wipo.int/pub/published_pct_sequences
 XX
 SQ Sequence 96 AA;
 Query Match 89.8%; Score 53; DB 4; Length 96;
 Best Local Similarity 90.9%; Pred. No. 20;
 Matches 10; Conservative 0; Mismatches 1; Indels 0; Gaps 0;
 Qy 1 CCCSSSSSSSS 11
 |||||
 Db 35 CCCSSSSSSSFS 45
 RESULT 7
 ID ABG28278 standard; protein; 103 AA.
 XX
 AC ABG28278;
 XX
 DT 18-FEB-2002 (first entry)
 XX
 DE Novel human diagnostic protein #28269.
 XX
 KW Human; chromosome mapping; gene mapping; gene therapy; forensic;
 KW food supplement; medical imaging; diagnostic; genetic disorder.
 XX
 OS Homo sapiens.
 XX
 PN WO200175067-A2.
 XX
 PD 11-OCT-2001.
 XX
 PF 30-MAR-2001; 2001WO-US008631.
 XX
 PR 31-MAR-2000; 2000US-00540217.
 PR 23-AUG-2000; 2000US-00649167.
 XX
 PA (HYSE-) HYSEQ INC.
 XX
 PI Drmanac RT, Liu C, Tang YT;
 XX
 DR WPI; 2001-639362/73.
 DR N-PSDB; AAS92465.
 XX
 PT New isolated polynucleotide and encoded polypeptides, useful in
 PT diagnostics, forensics, gene mapping, identification of mutations
 PT responsible for genetic disorders or other traits and to assess
 PT biodiversity.
 XX
 PS Claim 20; SEQ ID NO 58637; 103pp; English.
 XX
 CC The invention relates to isolated polynucleotide (I) and polypeptide (II)
 CC sequences. (I) is useful as hybridisation probes, polymerase chain
 CC reaction (PCR) primers, oligomers, and for chromosome and gene mapping,
 CC and in recombinant production of (II). The polynucleotides are also used
 CC in diagnostics as expressed sequence tags for identifying expressed
 CC genes. (I) is useful in gene therapy techniques to restore normal
 CC activity of (II) or to treat disease states involving (II). (II) is
 CC useful for generating antibodies against it, detecting or quantitating a
 CC polypeptide in tissue, as molecular weight markers and as a food
 CC supplement. (II) and its binding partners are useful in medical imaging
 CC of sites expressing (II). (I) and (II) are useful for treating disorders
 CC involving aberrant protein expression or biological activity. The
 CC polypeptide and polynucleotide sequences have applications in
 CC diagnostics, forensics, gene mapping, identification of mutations
 CC responsible for genetic disorders or other traits to assess biodiversity
 CC and to produce other types of data and products dependent on DNA and

CC amino acid sequences. ABG00010-ABG30377 represent novel human diagnostic
 CC amino acid sequences of the invention. Note: The sequence data for this
 CC patent did not appear in the printed specification, but was obtained in
 CC electronic format directly from WIPO at
 CC ftp.wipo.int/pub/published_pct_sequences
 XX
 SQ Sequence 103 AA;
 Query Match 86.4%; Score 51; DB 4; Length 103;
 Best Local Similarity 81.8%; Pred. No. 36;
 Matches 9; Conservative 1; Mismatches 1; Indels 0; Gaps 0;
 Qy 1 CCCSSSSSSSS 11
 |||||
 Db 57 CCCSSSHNSSS 67
 RESULT 8
 ID ADK34141
 ID ADK34141 standard; protein; 120 AA.
 XX
 AC ADK34141;
 XX
 DT 06-MAY-2004 (first entry)
 XX
 DE Novel human polypeptide SeqID6223.
 XX
 KW antiarthritic; antiparkinsonian; neuroprotective; nootropic;
 KW immunosuppressive; cytostatic; antipsoriatic; antiinflammatory;
 KW antibacterial; antiviral; antifungal; antiparasitic; gene therapy;
 KW arthritis; Parkinson's; Alzheimer's; autoimmune disease; cancer;
 KW psoriasis; inflammatory bowel disease; infection; bacteria; virus;
 KW fungus; parasite; human.
 XX
 OS Homo sapiens.
 XX
 FH Key Location/Qualifiers
 FT Misc-difference 1..120 /label= OTHER
 FT /note= "OTHER= All Xaa's in this sequence are unknown
 FT amino acids or the site of a stop codon within the DNA
 FT sequence"
 XX
 PN WO200216439-A2.
 XX
 PD 28-FEB-2002.
 XX
 PF 05-MAR-2001; 2001WO-US004941.
 XX
 PR 07-MAR-2000; 2000US-00519705.
 PR 19-MAY-2000; 2000US-00574454.
 XX
 PA (HYSE-) HYSEQ INC.
 XX
 PI Tang YT, Liu C, Drmanac RT;
 XX
 DR WPI; 2002-280918/32.
 XX
 PT Isolated polynucleotide encoding bone marrow derived polypeptides useful
 PT for treating, e.g., Parkinson's, Alzheimer's, cancer, arthritis, Crohn's
 PT disease, and inflammatory bowel disease.
 XX
 PS Claim 20; SEQ ID NO 6223; 504pp; English.
 XX
 CC This invention relates to a novel isolated polynucleotide comprising a
 CC nucleotide sequence selected from one of 1680 sequences, a mature protein
 CC coding portion of them, an active domain of them and their complementary
 CC sequences. The invention may be useful for the production of compounds
 CC with an antiarthritic, antiparkinsonian, neuroprotective, nootropic,
 CC immunosuppressive, cytostatic, antipsoriatic, antiinflammatory,
 CC antibacterial, antiviral, antifungal or antiparasitic activity. In
 CC addition, the disclosed sequences may be useful for gene therapy. The
 CC polypeptides or their antibodies are useful for treating many diseases

CC such as arthritis, Parkinson's, Alzheimer's, autoimmune diseases, cancer,
 CC psoriasis, inflammatory bowel disease and infections caused by bacteria,
 CC viruses, fungi or parasites. The present sequence is that of a human
 CC polypeptide of the invention.

XX SQ Sequence 120 AA;

Query Match 84.7%; Score 50; DB 5; Length 120;
 Best Local Similarity 90.9%; Pred. No. 53;
 Matches 10; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 1 CCCSSSSSSSS 11
 ||| |||||
 Db 32 CCASSSSSSSS 42

RESULT 9
 AAEE38440
 ID AAE38440 standard; protein; 940 AA.

XX AC AAE38440;

XX 20-NOV-2003 (first entry)

XX DE Human FISH protein.

XX Human; SH3 domain; Src homology 3 domain; proline-rich region; allergy;
 KW neuronal Wiskott-Aldrich syndrome protein; CD2 associated protein; PRR;
 KW epidermal growth factor receptor; Src-associated in mitosis; cortactin;
 KW Cbl adaptor protein; T-cell mediated disorder; autoimmune disease; FISH;
 KW transplant rejection; cancer; endophilin; gene therapy; N-WASP; asthma;
 KW ASAP1; Sam68; EGFR; CD2AP.

XX OS Homo sapiens.

XX PN WO2003068808-A1.

XX PD 21-AUG-2003.

XX PF 12-FEB-2003; 2003WO-AU000177.

XX PR 13-FEB-2002; 2002AU-00000514.

XX PR 08-AUG-2002; 2002AU-00950698.

XX (GARV-) GARVAN INST MEDICAL RES.

XX PI Daly R;

XX WPI; 2003-646476/61.

XX New isolated recombinant complex, useful for treating a disorder
 PT characterized by aberrant EGFR endocytosis, e.g. cancer, T-cell mediated
 PT disorders such as autoimmune diseases, transplant rejection, allergy or
 PT asthma.

PS Claim 19; Page 160-163; 181pp; English.

XX The present invention provides novel protein complexes involving SH3 (Src
 CC homology 3 domain) domain-containing proteins and proteins having one or
 CC more proline-rich regions (PRRs). Specifically the invention relates to
 CC the protein complexes comprising a cortactin protein and/or a CD2AP (CD2
 CC associated protein) protein or a portion thereof and one or more other
 CC proteins selected from the group consisting of Cbl adaptor protein,
 CC endophilin, ASAP1, N-WASP (neuronal Wiskott-Aldrich syndrome protein),
 CC FISH and Sam68 (Src-associated in mitosis). Protein complexes of the
 CC invention are useful for treating disorders characterised by aberrant
 CC epidermal growth factor (EGF) receptor (EGFR) endocytosis e.g. cancer, T-
 CC cell mediated disorders such as autoimmune diseases, asthma, transplant
 CC rejection and allergy. The invention is also useful in gene therapy. The
 CC present sequence is human FISH protein

XX SQ Sequence 940 AA;

Query Match 84.7%; Score 50; DB 6; Length 940;
 Best Local Similarity 100.0%; Pred. No. 2.9e+02;
 Matches 10; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 2 CCSSSSSSSS 11
 |||||
 Db 507 CCSSSSSSSS 516

RESULT 10

AAU14174

ID AAU14174 standard; protein; 968 AA.

XX AC AAU14174;

XX 24-OCT-2001 (first entry)

XX DE Human novel protein #45.

XX Human; novel protein; Antianaemic; osteopathic; antiinflammatory;
 KW immunomodulatory; cytostatic; neuroprotective; vulnary; neurotropic;
 KW anticonvulsant; antiarthritic; cerebroprotective; antifungal; antiviral;
 KW antibacterial; antiallergic; dermatological; haemostatic; antiaschmatic;
 KW thrombolytic; immunogen; antibody; gene therapy; neurological disorder;
 KW Parkinson's disease; inflammatory disorder; cancer; asthma; osteoporosis;
 KW tissue regeneration; immune disorder.

XX OS Homo sapiens.

XX PN WO200155437-A2.

XX PD 02-AUG-2001.

XX 25-JAN-2001; 2001WO-US002623.

XX 25-JAN-2000; 2000US-00491404.

XX (HYSE-) HYSEQ INC.

XX Tang YT, Liu C, Drmanac RT;

XX WPI; 2001-451939/48.

XX N-PSDB; AAS22479.

XX Isolated polypeptides useful for treating anti-inflammatory diseases,
 PT nervous system disorders, and for regenerating bone and cartilage.

XX Example 4; Page 551-552; 894pp; English.

XX The invention relates to polynucleotides encoding novel human proteins or
 CC their active domains. The polypeptides, polynucleotides and antibodies
 CC raised against the polypeptides are used in a method of treatment of a
 CC mammal and prevention of disorders caused by the aberrant protein
 CC expression or activity. The polypeptides can be used as molecular weight
 CC markers, food supplements, and in antibody production. The polypeptides
 CC are used to identify compounds which bind to the polypeptides.
 CC Polynucleotides of the invention are used as probes and primers, for
 CC sequencing, for chromosome or gene mapping, in the production of
 CC recombinant proteins, and in generating anti-sense DNA or RNA and in gene
 CC therapy. Polypeptides of the invention can be used to target drugs to a
 CC tumour, in assays to determine biological activity, to raise
 CC antibodies/elicite an immune response, to determine quantitative protein
 CC levels, as tissue markers, and to isolate receptors or ligands.

XX Polypeptides of the invention may also be useful in treating platelet
 CC disorders, stem cell disorders, regenerating bone, cartilage, tendon,
 CC ligament and/or nerve tissue, wound healing, treating burns, promoting
 CC the proliferation, differentiation and survival of stem cells, as a
 CC contraceptive, treating osteoporosis and osteoarthritis, anaemia,
 CC Alzheimer's, Parkinson's and Huntington's diseases, amyotrophic lateral
 CC sclerosis, stroke, immune deficiencies resulting from bacterial, viral or
 CC fungal infection or from autoimmunity, cancer, allergy, asthma, graft-
 CC versus-host disease, eczema, haemophilia, thrombosis, anti-inflammatory
 CC diseases, nervous system disorders, and infection. The present sequence

CC represents a protein of the invention

SQ Sequence 968 AA;

Query Match 84.7%; Score 50; DB 4; Length 968;

Best Local Similarity 100.0%; Pred. No. 2.9e+02;

Matches 10; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 2 CCSSSSSSSS 11

|||||

535 CCSSSSSSSS 544

RESULT 11

ADA55008
ID ADA55008 standard; protein; 1054 AA.

AC ADA55008;

XX 20-NOV-2003 (first entry)

XX Human protein, SEQ ID 2576.

XX Cytostatic; Anti-inflammatory; Osteopathic; Neuroprotective; Nootropic;
KW Gene Therapy; human; secretory protein; membrane proteins; cancer;
KW inflammatory disease; osteoporosis; neurological disease.

XX Homo sapiens.

XX EPI293569-A2.

XX 19-MAR-2003.

XX 21-MAR-2002; 2002EP-00006586.

XX 14-SEP-2001; 2001JP-00328381.

XX 24-JAN-2002; 2002US-0350435P.

XX (HELI-) HELIX RES INST.

XX (REAS-) RES ASSOC BIOTECHNOLOGY.

XX Isogai T, Sugiyama T, Otsuki T, Wakamatsu A, Sato H, Ishii S;

PI Yamamoto J, Isono Y, Hio Y, Otsuka K, Nagai K, Tamechika I;

PI Seki N, Yoshikawa T, Otsuka M, Nagahari K, Masuho Y;

XX WPI; 2003-395539/38.

XX N-ESDB; ADA53369.

XX New polynucleotides encoding full-length polypeptides, e.g. secretory
PT and/or membrane proteins, useful for developing medicines for diseases in
PT which the gene is involved, or as target molecules for gene therapy.

XX Claim 14; SEQ ID NO 2576; 205pp; English.

XX The present invention relates to novel human secretory or membrane
CC proteins (ADA54072-ADA55710) and their coding sequences (ADA52433-
CC ADA54071). The coding sequences are useful in the gene therapy of
CC diseases caused by abnormalities of the proteins, e.g. cancer,
CC inflammatory diseases, osteoporosis or neurological disease.

SQ Sequence 1054 AA;

Query Match 84.7%; Score 50; DB 6; Length 1054;

Best Local Similarity 100.0%; Pred. No. 3.1e+02;

Matches 10; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 2 CCSSSSSSSS 11

|||||

621 CCSSSSSSSS 630

RESULT 12

AAG41076

ID AAG41076 standard; protein; 91 AA.

XX AAG41076;

XX 18-OCT-2000 (first entry)

XX Zea mays protein fragment SEQ ID NO: 51058.

XX Protein identification; signal transduction pathway; metabolic pathway;
KW hybridisation assay; genetic mapping; gene expression control; promoter;
KW termination sequence; corn.

XX Zea mays subsp. mays.

XX EPI033405-A2.

XX 06-SEP-2000.

XX 25-FEB-2000; 2000EP-00301439.

XX 25-FEB-1999; 99US-0121825P.

XX 05-MAR-1999; 99US-0123180P.

XX 09-MAR-1999; 99US-0123548P.

XX 23-MAR-1999; 99US-0125788P.

XX 25-MAR-1999; 99US-0126264P.

XX 29-MAR-1999; 99US-0126785P.

XX 01-APR-1999; 99US-0127462P.

XX 06-APR-1999; 99US-0128234P.

XX 08-APR-1999; 99US-0128714P.

XX 16-APR-1999; 99US-0129845P.

XX 19-APR-1999; 99US-0130077P.

XX 21-APR-1999; 99US-0130449P.

XX 23-APR-1999; 99US-0130510P.

XX 28-APR-1999; 99US-0130891P.

XX 30-APR-1999; 99US-0131449P.

XX 30-APR-1999; 99US-0132048P.

XX 04-MAY-1999; 99US-0132407P.

XX 05-MAY-1999; 99US-0132484P.

XX 06-MAY-1999; 99US-0132485P.

XX 06-MAY-1999; 99US-0132486P.

XX 07-MAY-1999; 99US-0132487P.

XX 11-MAY-1999; 99US-0132863P.

XX 14-MAY-1999; 99US-0134256P.

XX 14-MAY-1999; 99US-0134218P.

XX 14-MAY-1999; 99US-0134219P.

XX 14-MAY-1999; 99US-0134221P.

XX 14-MAY-1999; 99US-0134370P.

XX 18-MAY-1999; 99US-0134768P.

XX 19-MAY-1999; 99US-0134941P.

XX 20-MAY-1999; 99US-0135124P.

XX 21-MAY-1999; 99US-0135353P.

XX 24-MAY-1999; 99US-0135629P.

XX 25-MAY-1999; 99US-0136021P.

XX 27-MAY-1999; 99US-0136392P.

XX 28-MAY-1999; 99US-0136782P.

XX 01-JUN-1999; 99US-0137222P.

XX 03-JUN-1999; 99US-0137528P.

XX 04-JUN-1999; 99US-0137502P.

XX 07-JUN-1999; 99US-0137724P.

XX 08-JUN-1999; 99US-0138094P.

XX 10-JUN-1999; 99US-0138540P.

XX 10-JUN-1999; 99US-0138847P.

XX 14-JUN-1999; 99US-0139119P.

XX 16-JUN-1999; 99US-0139452P.

XX 16-JUN-1999; 99US-0139453P.

XX 17-JUN-1999; 99US-0139492P.

XX 18-JUN-1999; 99US-0139454P.

XX 18-JUN-1999; 99US-0139455P.

XX 18-JUN-1999; 99US-0139456P.

XX 18-JUN-1999; 99US-0139457P.

XX 18-JUN-1999; 99US-0139458P.

XX 18-JUN-1999; 99US-0139459P.

XX 18-JUN-1999; 99US-0139460P.

```
PR 18-JUN-1999; 99US-0139461P.
PR 18-JUN-1999; 99US-0139462P.
PR 18-JUN-1999; 99US-0139463P.
PR 18-JUN-1999; 99US-0139750P.
PR 18-JUN-1999; 99US-0139763P.
PR 21-JUN-1999; 99US-0139817P.
PR 22-JUN-1999; 99US-0139899P.
PR 23-JUN-1999; 99US-0140353P.
PR 23-JUN-1999; 99US-0140354P.
PR 23-JUN-1999; 99US-0140695P.
PR 28-JUN-1999; 99US-0140823P.
PR 29-JUN-1999; 99US-0140991P.
PR 30-JUN-1999; 99US-0141287P.
PR 01-JUL-1999; 99US-0141842P.
PR 01-JUL-1999; 99US-0142154P.
PR 02-JUL-1999; 99US-0142055P.
PR 06-JUL-1999; 99US-0142390P.
PR 08-JUL-1999; 99US-0142803P.
PR 09-JUL-1999; 99US-0142920P.
PR 12-JUL-1999; 99US-0142977P.
PR 13-JUL-1999; 99US-0143542P.
PR 14-JUL-1999; 99US-0143624P.
PR 15-JUL-1999; 99US-0144005P.
PR 16-JUL-1999; 99US-0144085P.
PR 16-JUL-1999; 99US-0144086P.
PR 19-JUL-1999; 99US-0144325P.
PR 19-JUL-1999; 99US-0144331P.
PR 19-JUL-1999; 99US-0144332P.
PR 19-JUL-1999; 99US-0144333P.
PR 19-JUL-1999; 99US-0144334P.
PR 19-JUL-1999; 99US-0144335P.
PR 20-JUL-1999; 99US-0144352P.
PR 20-JUL-1999; 99US-0144632P.
PR 20-JUL-1999; 99US-0144884P.
PR 21-JUL-1999; 99US-0144814P.
PR 21-JUL-1999; 99US-0145086P.
PR 21-JUL-1999; 99US-0145088P.
PR 22-JUL-1999; 99US-0145085P.
PR 22-JUL-1999; 99US-0145087P.
PR 22-JUL-1999; 99US-0145089P.
PR 22-JUL-1999; 99US-0145192P.
PR 23-JUL-1999; 99US-0145145P.
PR 23-JUL-1999; 99US-0145218P.
PR 23-JUL-1999; 99US-0145224P.
PR 26-JUL-1999; 99US-0145276P.
PR 27-JUL-1999; 99US-0145913P.
PR 27-JUL-1999; 99US-0145918P.
PR 27-JUL-1999; 99US-0145919P.
PR 28-JUL-1999; 99US-0145951P.
PR 02-AUG-1999; 99US-0146386P.
PR 02-AUG-1999; 99US-0146388P.
PR 02-AUG-1999; 99US-0146389P.
PR 03-AUG-1999; 99US-0147038P.
PR 04-AUG-1999; 99US-0147204P.
PR 04-AUG-1999; 99US-0147302P.
PR 05-AUG-1999; 99US-0147192P.
PR 05-AUG-1999; 99US-0147260P.
PR 06-AUG-1999; 99US-0147303P.
PR 06-AUG-1999; 99US-0147416P.
PR 09-AUG-1999; 99US-0147493P.
PR 09-AUG-1999; 99US-0147935P.
PR 10-AUG-1999; 99US-0148171P.
PR 11-AUG-1999; 99US-0148319P.
PR 12-AUG-1999; 99US-0148341P.
PR 13-AUG-1999; 99US-0148565P.
PR 13-AUG-1999; 99US-0148684P.
PR 16-AUG-1999; 99US-0149368P.
PR 17-AUG-1999; 99US-0149175P.
PR 18-AUG-1999; 99US-0149426P.
PR 20-AUG-1999; 99US-0149722P.
PR 20-AUG-1999; 99US-0149723P.
PR 20-AUG-1999; 99US-0149929P.
PR 23-AUG-1999; 99US-0149902P.
```

```
PR 23-AUG-1999; 99US-0149930P.
PR 25-AUG-1999; 99US-0150566P.
PR 26-AUG-1999; 99US-0150884P.
PR 27-AUG-1999; 99US-0151065P.
PR 27-AUG-1999; 99US-0151066P.
PR 27-AUG-1999; 99US-0151080P.
PR 30-AUG-1999; 99US-0151303P.
PR 31-AUG-1999; 99US-0151438P.
PR 01-SEP-1999; 99US-0151930P.
PR 07-SEP-1999; 99US-0152363P.
PR 10-SEP-1999; 99US-0153070P.
PR 13-SEP-1999; 99US-0153758P.
PR 15-SEP-1999; 99US-0154018P.
PR 16-SEP-1999; 99US-0154039P.
PR 20-SEP-1999; 99US-0154779P.
PR 22-SEP-1999; 99US-0155139P.
PR 23-SEP-1999; 99US-0155486P.
PR 24-SEP-1999; 99US-0155659P.
PR 28-SEP-1999; 99US-0156458P.
PR 29-SEP-1999; 99US-0156596P.
PR 04-OCT-1999; 99US-0157117P.
PR 05-OCT-1999; 99US-0157753P.
PR 06-OCT-1999; 99US-0157865P.
PR 07-OCT-1999; 99US-0158029P.
PR 08-OCT-1999; 99US-0158232P.
PR 12-OCT-1999; 99US-0158369P.
PR 13-OCT-1999; 99US-0159293P.
PR 13-OCT-1999; 99US-0159294P.
PR 13-OCT-1999; 99US-0159295P.
PR 14-OCT-1999; 99US-0159330P.
PR 14-OCT-1999; 99US-0159331P.
PR 14-OCT-1999; 99US-0159637P.
PR 14-OCT-1999; 99US-0159638P.
PR 18-OCT-1999; 99US-0159584P.
PR 21-OCT-1999; 99US-0160741P.
PR 21-OCT-1999; 99US-0160767P.
PR 21-OCT-1999; 99US-0160768P.
PR 21-OCT-1999; 99US-0160770P.
PR 21-OCT-1999; 99US-0160814P.
PR 21-OCT-1999; 99US-0160815P.
PR 22-OCT-1999; 99US-0160980P.
PR 22-OCT-1999; 99US-0160981P.
PR 22-OCT-1999; 99US-0160989P.
PR 25-OCT-1999; 99US-0161404P.
PR 25-OCT-1999; 99US-0161405P.
PR 25-OCT-1999; 99US-0161406P.
PR 26-OCT-1999; 99US-0161359P.
PR 26-OCT-1999; 99US-0161360P.
PR 26-OCT-1999; 99US-0161361P.
PR 28-OCT-1999; 99US-0161920P.
PR 28-OCT-1999; 99US-0161992P.
PR 28-OCT-1999; 99US-0161993P.
PR 29-OCT-1999; 99US-0162142P.
```

Query Match 81.4%; Score 48; DB 3; Length 91;
Best Local Similarity 72.7%; Pred. No. 73;
Matches 8; Conservative 2; Mismatches 1; Indels 0; Gaps 0;

QY 1 CCCSSSSSSSS 11
||| :|:
Db 67 CCCSRTSASS 77

RESULT 13

AAU18125
ID AAU18125 standard; protein; 109 AA.
XX
AC AAU18125;
XX
DT 21-NOV-2001 (first entry)
XX
DE Novel human uterine motility-association polypeptide #32.

XX Human; uterine motility-association disorder; uterus; pregnancy; labour;
KW menstrual cycle; gene therapy.
XX Homo sapiens.
XX W0200155201-A1.
XX 02-AUG-2001.
XX 17-JAN-2001; 2001WO-US0011317.
XX 31-JAN-2000; 2000US-0179065P.
PR 04-FEB-2000; 2000US-0180628P.
PR 24-FEB-2000; 2000US-0184664P.
PR 02-MAR-2000; 2000US-0186350P.
PR 16-MAR-2000; 2000US-0189874P.
PR 17-MAR-2000; 2000US-0190076P.
PR 18-APR-2000; 2000US-0198123P.
PR 19-MAY-2000; 2000US-0205515P.
PR 07-JUN-2000; 2000US-0209467P.
PR 28-JUN-2000; 2000US-0214886P.
PR 30-JUN-2000; 2000US-0215135P.
PR 01-JUL-2000; 2000US-0216647P.
PR 07-JUL-2000; 2000US-0216880P.
PR 11-JUL-2000; 2000US-0217496P.
PR 14-JUL-2000; 2000US-0218290P.
PR 26-JUL-2000; 2000US-0220963P.
PR 26-JUL-2000; 2000US-0220964P.
PR 14-AUG-2000; 2000US-0224518P.
PR 14-AUG-2000; 2000US-0225213P.
PR 14-AUG-2000; 2000US-0225447P.
PR 14-AUG-2000; 2000US-0225266P.
PR 14-AUG-2000; 2000US-0225267P.
PR 14-AUG-2000; 2000US-0225268P.
PR 14-AUG-2000; 2000US-0225270P.
PR 14-AUG-2000; 2000US-0225447P.
PR 14-AUG-2000; 2000US-0225757P.
PR 14-AUG-2000; 2000US-0225758P.
PR 14-AUG-2000; 2000US-0225759P.
PR 18-AUG-2000; 2000US-0226279P.
PR 22-AUG-2000; 2000US-0226681P.
PR 22-AUG-2000; 2000US-0226868P.
PR 22-AUG-2000; 2000US-0227182P.
PR 23-AUG-2000; 2000US-0227009P.
PR 30-AUG-2000; 2000US-0228924P.
PR 01-SEP-2000; 2000US-0229287P.
PR 01-SEP-2000; 2000US-0229343P.
PR 01-SEP-2000; 2000US-0229344P.
PR 01-SEP-2000; 2000US-0229345P.
PR 05-SEP-2000; 2000US-0229509P.
PR 05-SEP-2000; 2000US-0229513P.
PR 06-SEP-2000; 2000US-0230437P.
PR 08-SEP-2000; 2000US-0230438P.
PR 08-SEP-2000; 2000US-0231242P.
PR 08-SEP-2000; 2000US-0231243P.
PR 08-SEP-2000; 2000US-0231244P.
PR 08-SEP-2000; 2000US-0231413P.
PR 08-SEP-2000; 2000US-0231414P.
PR 08-SEP-2000; 2000US-0232080P.
PR 12-SEP-2000; 2000US-0232081P.
PR 14-SEP-2000; 2000US-0231968P.
PR 14-SEP-2000; 2000US-0232398P.
PR 14-SEP-2000; 2000US-0232399P.
PR 14-SEP-2000; 2000US-0232400P.
PR 14-SEP-2000; 2000US-0232401P.
PR 14-SEP-2000; 2000US-0233063P.
PR 14-SEP-2000; 2000US-0233064P.
PR 14-SEP-2000; 2000US-0233065P.
PR 21-SEP-2000; 2000US-0234223P.
PR 21-SEP-2000; 2000US-0234274P.
PR 25-SEP-2000; 2000US-0234997P.
PR 26-SEP-2000; 2000US-0235484P.
PR 27-SEP-2000; 2000US-0235834P.
PR 27-SEP-2000; 2000US-0235836P.
PR 29-SEP-2000; 2000US-0236327P.
PR 29-SEP-2000; 2000US-0236367P.
PR 29-SEP-2000; 2000US-0236368P.
PR 29-SEP-2000; 2000US-0236369P.
PR 02-OCT-2000; 2000US-0236370P.
PR 02-OCT-2000; 2000US-0236802P.
PR 02-OCT-2000; 2000US-0237037P.
PR 02-OCT-2000; 2000US-0237038P.
PR 02-OCT-2000; 2000US-0237039P.
PR 02-OCT-2000; 2000US-0237040P.
PR 13-OCT-2000; 2000US-0239355P.
PR 13-OCT-2000; 2000US-0239377P.
PR 20-OCT-2000; 2000US-0240960P.
PR 20-OCT-2000; 2000US-0241221P.
PR 20-OCT-2000; 2000US-0241785P.
PR 20-OCT-2000; 2000US-0241786P.
PR 20-OCT-2000; 2000US-0241787P.
PR 20-OCT-2000; 2000US-0241808P.
PR 20-OCT-2000; 2000US-0241809P.
PR 20-OCT-2000; 2000US-0241826P.
PR 01-NOV-2000; 2000US-0244617P.
PR 08-NOV-2000; 2000US-0246474P.
PR 08-NOV-2000; 2000US-0246475P.
PR 08-NOV-2000; 2000US-0246476P.
PR 08-NOV-2000; 2000US-0246477P.
PR 08-NOV-2000; 2000US-0246478P.
PR 08-NOV-2000; 2000US-0246523P.
PR 08-NOV-2000; 2000US-0246524P.
PR 08-NOV-2000; 2000US-0246525P.
PR 08-NOV-2000; 2000US-0246526P.
PR 08-NOV-2000; 2000US-0246527P.
PR 08-NOV-2000; 2000US-0246528P.
PR 08-NOV-2000; 2000US-0246532P.
PR 08-NOV-2000; 2000US-0246609P.
PR 08-NOV-2000; 2000US-0246610P.
PR 08-NOV-2000; 2000US-0246611P.
PR 08-NOV-2000; 2000US-0246613P.
PR 17-NOV-2000; 2000US-0249207P.
PR 17-NOV-2000; 2000US-0249208P.
PR 17-NOV-2000; 2000US-0249209P.
PR 17-NOV-2000; 2000US-0249210P.
PR 17-NOV-2000; 2000US-0249211P.
PR 17-NOV-2000; 2000US-0249212P.
PR 17-NOV-2000; 2000US-0249213P.
PR 17-NOV-2000; 2000US-0249214P.
PR 17-NOV-2000; 2000US-0249215P.
PR 17-NOV-2000; 2000US-0249216P.
PR 17-NOV-2000; 2000US-0249217P.
PR 17-NOV-2000; 2000US-0249218P.
PR 17-NOV-2000; 2000US-0249244P.
PR 17-NOV-2000; 2000US-0249245P.
PR 17-NOV-2000; 2000US-0249264P.
PR 17-NOV-2000; 2000US-0249265P.
PR 17-NOV-2000; 2000US-0249277P.
PR 17-NOV-2000; 2000US-0249299P.
PR 01-DEC-2000; 2000US-0249300P.
PR 01-DEC-2000; 2000US-0250391P.
PR 05-DEC-2000; 2000US-0251030P.
PR 05-DEC-2000; 2000US-0251988P.
PR 06-DEC-2000; 2000US-0256719P.
PR 08-DEC-2000; 2000US-0251479P.
PR 08-DEC-2000; 2000US-0251856P.
PR 08-DEC-2000; 2000US-0251868P.
PR 08-DEC-2000; 2000US-0251869P.
PR 08-DEC-2000; 2000US-0251989P.
PR 08-DEC-2000; 2000US-0251990P.

PR	18-APR-2000;	2000US-01981233P
PR	19-MAY-2000;	2000US-02005155P
PR	17-JUN-2000;	2000US-0209467P
PR	28-JUN-2000;	2000US-0214886P
PR	30-JUN-2000;	2000US-0215135P
PR	07-JUL-2000;	2000US-0216647P
PR	07-JUL-2000;	2000US-0216880P
PR	11-JUL-2000;	2000US-0217487P
PR	11-JUL-2000;	2000US-0217496P
PR	14-JUL-2000;	2000US-0218290P
PR	26-JUL-2000;	2000US-02205963P
PR	26-JUL-2000;	2000US-0220964P
PR	14-AUG-2000;	2000US-0224518P
PR	14-AUG-2000;	2000US-0224519P
PR	14-AUG-2000;	2000US-0225213P
PR	14-AUG-2000;	2000US-0225214P
PR	14-AUG-2000;	2000US-0225266P
PR	14-AUG-2000;	2000US-0225677P
PR	14-AUG-2000;	2000US-02252688P
PR	14-AUG-2000;	2000US-0225270P
PR	22-AUG-2000;	2000US-0226681P
PR	14-AUG-2000;	2000US-0225447P
PR	14-AUG-2000;	2000US-0225577P
PR	23-AUG-2000;	2000US-0221782P
PR	23-AUG-2000;	2000US-0227009P
PR	30-AUG-2000;	2000US-0228924P
PR	01-SEP-2000;	2000US-0229287P
PR	01-SEP-2000;	2000US-0229343P
PR	01-SEP-2000;	2000US-0229344P
PR	01-SEP-2000;	2000US-0229345P
PR	05-SEP-2000;	2000US-0229509P
PR	03-SEP-2000;	2000US-0229513P
PR	06-SEP-2000;	2000US-0230437P
PR	08-SEP-2000;	2000US-0230438P
PR	08-SEP-2000;	2000US-0231242P
PR	08-SEP-2000;	2000US-0231243P
PR	08-SEP-2000;	2000US-0231244P
PR	08-SEP-2000;	2000US-0233413P
PR	08-SEP-2000;	2000US-0233414P
PR	08-SEP-2000;	2000US-02332080P
PR	08-SEP-2000;	2000US-0232081P
PR	12-SEP-2000;	2000US-0231968P
PR	14-SEP-2000;	2000US-0233397P
PR	14-SEP-2000;	2000US-0233398P
PR	14-SEP-2000;	2000US-0233399P
PR	21-SEP-2000;	2000US-0234223P
PR	21-SEP-2000;	2000US-0234223P
PR	14-SEP-2000;	2000US-0233400P
PR	14-SEP-2000;	2000US-0233401P
PR	14-SEP-2000;	2000US-0233063P
PR	14-SEP-2000;	2000US-0233064P
PR	14-SEP-2000;	2000US-0233065P
PR	21-SEP-2000;	2000US-0234223P
PR	21-SEP-2000;	2000US-0234223P
PR	25-SEP-2000;	2000US-0234997P
PR	25-SEP-2000;	2000US-0234988P
PR	26-SEP-2000;	2000US-0233484P
PR	27-SEP-2000;	2000US-0233634P
PR	27-SEP-2000;	2000US-0233636P
PR	29-SEP-2000;	2000US-0236327P
PR	29-SEP-2000;	2000US-0236367P
PR	29-SEP-2000;	2000US-0233688P
PR	29-SEP-2000;	2000US-0233699P
PR	29-SEP-2000;	2000US-0236370P
PR	02-OCT-2000;	2000US-02336802P
PR	02-OCT-2000;	2000US-0237037P
PR	02-OCT-2000;	2000US-0237038P
PR	02-OCT-2000;	2000US-0237039P
PR	02-OCT-2000;	2000US-0237040P
PR	13-OCT-2000;	2000US-0239935P
PR	13-OCT-2000;	2000US-0239937P

Qy 2 C C S S S S S S S S 1 1
 | | | | | | | | | |
 91 C C S A S S S S S S 1 0 0
 | | | | | | | | | |
 Db

08-SEP-2000;	2000US-02312433P;	PR
08-SEP-2000;	2000US-02312443P;	PR
08-SEP-2000;	2000US-02313413P;	PR
08-SEP-2000;	2000US-0231414P;	PR
08-SEP-2000;	2000US-02320800P;	PR
08-SEP-2000;	2000US-0232081P;	PR
12-SEP-2000;	2000US-0231968P;	PR
14-SEP-2000;	2000US-0233397P;	PR
14-SEP-2000;	2000US-0233398P;	PR
14-SEP-2000;	2000US-0233399P;	PR
14-SEP-2000;	2000US-0233400P;	PR
14-SEP-2000;	2000US-0233401P;	PR
14-SEP-2000;	2000US-0233063P;	PR
14-SEP-2000;	2000US-0233064P;	PR
14-SEP-2000;	2000US-0233065P;	PR
21-SEP-2000;	2000US-0234223P;	PR
21-SEP-2000;	2000US-0234274P;	PR
25-SEP-2000;	2000US-0234997P;	PR
25-SEP-2000;	2000US-0234998P;	PR

PR	27-SEP-2000;	2000US-0233634P;
PR	27-SEP-2000;	2000US-0233634P;
PR	27-SEP-2000;	2000US-0233634P;
PR	27-SEP-2000;	2000US-0233636P;
PR	29-SEP-2000;	2000US-0233637P;
PR	29-SEP-2000;	2000US-0233637P;
PR	29-SEP-2000;	2000US-0233637P;
PR	29-SEP-2000;	2000US-0233638P;
PR	29-SEP-2000;	2000US-0233639P;
PR	29-SEP-2000;	2000US-0233639P;
PR	29-SEP-2000;	2000US-02336370P;
PR	02-OCT-2000;	2000US-02336802P;
PR	02-OCT-2000;	2000US-02337037P;
PR	02-OCT-2000;	2000US-02337038P;
PR	02-OCT-2000;	2000US-02337038P;
PR	02-OCT-2000;	2000US-02337039P;
PR	02-OCT-2000;	2000US-02337040P;
PR	13-OCT-2000;	2000US-02339357P;
PR	13-OCT-2000;	2000US-02339357P;
PR	13-OCT-2000;	2000US-02339357P;

PR 20-OCT-2000; 2000US-0240960P.
PR 20-OCT-2000; 2000US-0241321P.
PR 20-OCT-2000; 2000US-0241785P.
PR 20-OCT-2000; 2000US-0241786P.
PR 20-OCT-2000; 2000US-0241787P.
PR 20-OCT-2000; 2000US-0241808P.
PR 20-OCT-2000; 2000US-0241809P.
PR 20-OCT-2000; 2000US-0241826P.
PR 01-NOV-2000; 2000US-0244617P.
PR 08-NOV-2000; 2000US-0246474P.
PR 08-NOV-2000; 2000US-0246475P.
PR 08-NOV-2000; 2000US-0246476P.
PR 08-NOV-2000; 2000US-0246477P.
PR 08-NOV-2000; 2000US-0246478P.
PR 08-NOV-2000; 2000US-0246523P.
PR 08-NOV-2000; 2000US-0246524P.
PR 08-NOV-2000; 2000US-0246525P.
PR 08-NOV-2000; 2000US-0246526P.
PR 08-NOV-2000; 2000US-0246527P.
PR 08-NOV-2000; 2000US-0246528P.
PR 08-NOV-2000; 2000US-0246532P.
PR 08-NOV-2000; 2000US-0246609P.
PR 08-NOV-2000; 2000US-0246610P.
PR 08-NOV-2000; 2000US-0246611P.
PR 08-NOV-2000; 2000US-0246613P.
PR 17-NOV-2000; 2000US-0249207P.
PR 17-NOV-2000; 2000US-0249208P.
PR 17-NOV-2000; 2000US-0249209P.
PR 17-NOV-2000; 2000US-0249210P.
PR 17-NOV-2000; 2000US-0249211P.
PR 17-NOV-2000; 2000US-0249212P.
PR 17-NOV-2000; 2000US-0249213P.
PR 17-NOV-2000; 2000US-0249214P.
PR 17-NOV-2000; 2000US-0249215P.
PR 17-NOV-2000; 2000US-0249216P.
PR 17-NOV-2000; 2000US-0249217P.
PR 17-NOV-2000; 2000US-0249218P.
PR 17-NOV-2000; 2000US-0249244P.
PR 17-NOV-2000; 2000US-0249245P.
PR 17-NOV-2000; 2000US-0249246P.
PR 17-NOV-2000; 2000US-0249247P.
PR 17-NOV-2000; 2000US-0249255P.
PR 17-NOV-2000; 2000US-0249257P.
PR 17-NOV-2000; 2000US-0249299P.
PR 17-NOV-2000; 2000US-0249300P.
PR 01-DEC-2000; 2000US-0250160P.
PR 01-DEC-2000; 2000US-0250391P.
PR 05-DEC-2000; 2000US-0251030P.
PR 05-DEC-2000; 2000US-0251988P.
PR 05-DEC-2000; 2000US-0256719P.
PR 06-DEC-2000; 2000US-0251479P.
PR 08-DEC-2000; 2000US-0251856P.
PR 08-DEC-2000; 2000US-0251868P.
PR 08-DEC-2000; 2000US-0251869P.
PR 08-DEC-2000; 2000US-0251989P.
PR 08-DEC-2000; 2000US-0251990P.
PR 11-DEC-2000; 2000US-0254097P.
PR 05-JAN-2001; 2001US-0259678P.
XX
XX
PA (HUMA-) HUMAN GENOME SCI INC.
XX
XX Rosen CA, Barash SC, Ruben SM;
PI
XX
XX
DR WPI: 2001-488787/53.
DR N-PSDB; AAS30196.
XX
XX New polynucleotides and polypeptides, useful for diagnosing, treating,
PT preventing or prognosing e.g. kidney, cardiovascular, blood, electrolyte
PT imbalance or neoplastic disorders, autoimmune diseases, cancers.
XX
XX Claim 1; SEQ ID NO 114; 506pp; English.
PS
XX The invention relates to novel nucleic acids and polypeptides useful for
CC diagnosing, treating, preventing and/or prognosing disorders related to
CC

CC these polypeptides. The polynucleotides are especially useful in the
CC diagnosis, prognosis, prevention and/or treatment of diseases which
CC include kidney disorders (e.g. renal failure or nephritis),
CC cardiovascular disorders (e.g. hypertension or myocardial infarction),
CC blood disorders (e.g. anaemia or blood coagulation disorders),
CC electrolyte imbalance disorders (e.g. hyponatraemia or hyperkalaemia),
CC neoplastic disorders (e.g. nephroma or renal cell cancer), autoimmune
CC diseases, cancers, inflammatory diseases, reproductive system disorders,
CC endocrine disorders, neural activity and neurological disorders, wound
CC healing and respiratory disorders. AAU18644-AAU18715 represent the novel
CC human renal and cardiovascular-associated amino acid sequences of the
CC invention. Note: the sequence data for this patent did not form part of
CC the printed specification, but was obtained in electronic format directly
CC from WIPO at: ftp.wipo.int/pub/published_pct_sequences
XX
SQ Sequence 109 AA;

Query Match 79.7%; Score 47; DB 4; Length 109;
Best Local Similarity 90.0%; Pred. NO. 1.1e+02;
Matches 9; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

Qy 2 CCSSSSSSSS 11
Db 91 CCSASSSSSS 100
||||:|||||

RESULT 15
AB10317
ID AB10317 standard; protein; 109 AA.
XX
AC AB10317;
XX
DT 10-JAN-2002 (first entry)
XX
DE Human cDNA SEQ ID NO: 625.
XX
KW Human; gene therapy; neural disorder; immune system disorder;
KW muscular disorder; reproductive disorder; gastrointestinal disorder;
KW pulmonary disorder; cardiovascular disorder; renal disorder;
KW proliferative disorder; inflammation.
XX
OS Homo sapiens.
XX
XX WO200154474-A2.
PN
XX
PD 02-AUG-2001.
XX
PF 17-JAN-2001; 2001WO-US001349.
XX
XX 31-JAN-2000; 2000US-0179065P.
PR 04-FEB-2000; 2000US-0180628P.
PR 24-FEB-2000; 2000US-0184664P.
PR 02-MAR-2000; 2000US-0186350P.
PR 16-MAR-2000; 2000US-0189874P.
PR 17-MAR-2000; 2000US-0190076P.
PR 18-APR-2000; 2000US-0198123P.
PR 19-MAY-2000; 2000US-0205515P.
PR 07-JUN-2000; 2000US-0209467P.
PR 28-JUN-2000; 2000US-0214886P.
PR 30-JUN-2000; 2000US-0215135P.
PR 07-JUL-2000; 2000US-0216647P.
PR 07-JUL-2000; 2000US-0216880P.
PR 11-JUL-2000; 2000US-0217487P.
PR 11-JUL-2000; 2000US-0217496P.
PR 14-JUL-2000; 2000US-0218290P.
PR 25-JUL-2000; 2000US-0220963P.
PR 26-JUL-2000; 2000US-0220964P.
PR 14-AUG-2000; 2000US-0224518P.
PR 14-AUG-2000; 2000US-0224519P.
PR 14-AUG-2000; 2000US-0225213P.
PR 14-AUG-2000; 2000US-0225214P.
PR 14-AUG-2000; 2000US-0225266P.
PR 14-AUG-2000; 2000US-0225267P.

This Page Blank (uspto)

GenCore version 5.1.6
Copyright (c) 1993 - 2005 CompuGen Ltd.

OM protein - protein search, using sw model

Run on: November 15, 2005, 12:11:54 ; Search time 9.8421 Seconds
(without alignments)
107.536 Million cell updates/sec

Title: US-10-820-998-3
Perfect score: 59
Sequence: 1 CCCSSSSSSSS 11

Scoring table: BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 283416 seqs, 96216763 residues
Total number of hits satisfying chosen parameters: 283416

Minimum DB seq length: 0
Maximum DB seq length: 2000000000
Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : PIR 79:.*
1: pir1:.*
2: pir2:.*
3: pir3:.*
4: pir4:.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	56	94.9	268	2 A84746	probable NAM (no a
2	50	84.7	940	2 T00056	hypothetical prote
3	47	79.7	593	2 T43323	cell division cycl
4	47	79.7	1307	2 G96711	unknown protein, 9
5	45	76.3	113	2 H84725	hypothetical prote
6	44	74.6	99	2 D81551	hypothetical prote
7	44	74.6	268	2 T43486	hypothetical prote
8	44	74.6	747	2 T34329	hypothetical prote
9	43	72.9	294	2 T23682	hypothetical prote
10	43	72.9	413	1 S03831	homeotic protein S
11	43	72.9	546	2 S47300	gene F protein - r
12	43	72.9	596	2 I38228	Shb protein - huma
13	42	71.2	129	2 T00282	hypothetical prote
14	42	71.2	169	1 S18946	ultra high-sulfur
15	42	71.2	1205	2 T41887	hypothetical prote
16	41	69.5	184	2 T15525	hypothetical prote
17	41	69.5	389	2 T23167	hypothetical prote
18	41	69.5	437	2 S55392	msh protein - fruit
19	41	69.5	568	2 T05218	hypothetical prote
20	41	69.5	609	2 S57378	hypothetical prote
21	41	69.5	614	2 T09902	hypothetical prote
22	41	69.5	837	2 T48407	hypothetical prote
23	41	69.5	1520	2 T00273	hypothetical prote
24	41	69.5	2531	2 T16743	hypothetical prote
25	40.5	68.6	472	2 S51768	glycerol-3-phospha
26	40	67.8	344	2 T00695	hypothetical prote
27	40	67.8	402	2 JC5847	3-ketoacyl-acyl ca
28	40	67.8	453	2 T48240	hypothetical prote
29	40	67.8	585	2 T51361	nitrate transporte

nitrate transporte
transcription fact
transcription acti
hypothetical prote
transcription fact
hypothetical prote
hypothetical prote
probable membrane
hypothetical prote
flagellar basal bo
G protein-coupled
kinesin heavy chai
hypothetical prote
unknown protein Fl
envelope protein -
protein K04H4.2b (

30 40 67.8 600 2 G96720
31 40 67.8 688 2 T09941
32 40 67.8 750 2 T10864
33 40 67.8 750 2 T04980
34 40 67.8 790 2 T12203
35 39 66.1 101 2 A71007
36 39 66.1 128 2 T15101
37 39 66.1 160 2 S45981
38 39 66.1 233 2 T00566
39 39 66.1 262 2 S27304
40 39 66.1 330 2 A55689
41 39 66.1 439 2 T49189
42 39 66.1 464 2 T15161
43 39 66.1 509 2 D96794
44 39 66.1 669 2 A46511
45 39 66.1 739 2 B88553

ALIGNMENTS

RESULT 1
A84746
probable NAM (no apical meristem)-like protein [imported] - Arabidopsis thaliana
C:Species: Arabidopsis thaliana (mouse-ear cress)
C>Date: 02-Feb-2001 #sequence_revision 02-Feb-2001 #text_change 09-Jul-2004
C:Accession: A84746
R:Lin, X.; Kaul, S.; Rounsley, S.D.; Shea, T.P.; Benito, M.I.; Town, C.D.; Fujii, C.Y.; N
M.; Koo, H.; Moffat, K.S.; Cronin, L.A.; Shen, M.; VanAken, S.E.; Umayam, L.; Tallon, L.
euss, D.; Niernman, W.C.; White, O.; Eisen, J.A.; Salzberg, S.L.; Fraser, C.M.; Venter, J
Nature 402, 761-768, 1999
A:Title: Sequence and analysis of chromosome 2 of the plant Arabidopsis thaliana.
A:Reference number: A84420; MUID:20083487; PMID:10617197
A:Accession: A84746
A>Status: preliminary
A:Molecule type: DNA
A:Residues: 1-268 <STO>
A:Cross-references: UNIPROT:O22798; GB:AE002093; NID:g2459430; PIDN:AAB80665.1; GSPDB:GN
C:Genetics:
A:Gene: At2g33480
A:Map position: 2

Query Match 94.9%; Score 56; DB 2; Length 268;
Best Local Similarity 90.9%; Pred. No. 1.8;
Matches 10; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

Qy 1 CCCSSSSSSSS 11
Db 229 CCCSSSSSSSS 239

RESULT 2
T00056
hypothetical protein KIAA0418 - human
C:Species: Homo sapiens (man)
C>Date: 22-Jan-1999 #sequence_revision 22-Jan-1999 #text_change 09-Jul-2004
C:Accession: T00056
R:Teikawa, K.; Nagase, T.; Nakajima, D.; Seki, N.; Ohira, M.; Miyajima, N.; Tanaka, A.;
submitted to the EMBL Data Library, October 1997
A:Description: Prediction of the coding sequences of unidentified human genes. VIII. The
A:Reference number: Z14080
A:Accession: T00056
A>Status: preliminary; translated from GB/EMBL/DDBJ
A:Molecule type: mRNA
A:Residues: 1-940 <ISH>
A:Cross-references: UNIPROT:O43302; EMBL:AB007878; NID:d1179754; PIDN:BAA24848.1; PID:d1
C:Genetics:
A:Note: KIAA0418

Query Match 84.7%; Score 50; DB 2; Length 940;
Best Local Similarity 100.0%; Pred. No. 22;
Matches 10; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

```

OY      2 CCSSSSSSSS 11
      |||||
Db      507 CCSSSSSSSS 516
      |||||

RESULT 3
T43323
cell division cycle protein 23 - fission yeast (Schizosaccharomyces pombe)
C;Species: Schizosaccharomyces pombe
C;Date: 11-Jan-2000 #sequence_revision 11-Jan-2000 #text_change 09-Jul-2004
C;Accession: T43323; T39398
R;Kawasaki, Y.; Tanaka, H.; Murakami, H.; Okayama, H.; Tye, B.K.
submitted to the EMBL Data Library, February 1998
A;Description: The pre-replication complex impedes the progression of replication fork
A;Reference number: 222426
A;Accession: T43323
A;Status: preliminary; translated from GB/EMBL/DBJ
A;Molecule type: DNA
A;Residues: 1-593 <KAW>
A;Cross-references: UNIPROT:O42709; EMBL:AB011244; PIDN:BAA24935.1
R;Wood, V.; Rajandream, M.A.; Barrrell, B.G.; Rieger, M.
submitted to the EMBL Data Library, February 1999
A;Reference number: 221851
A;Accession: T39398
A;Status: preliminary; translated from GB/EMBL/DBJ
A;Molecule type: DNA
A;Residues: 1-593 <WOO>
A;Cross-references: EMBL:AL035548; PIDN:CAB37441.1; GSPDB:GNO0067; SPDB:SPBC1347.10
A;Experimental source: strain 972h-; cosmid c1347
C;Genetics:
A;Gene: cdc23; SPBC1347.10
A;Map position: 2

Query Match      79.7%; Score 47; DB 2; Length 593;
Best Local Similarity 81.8%; Pred. No. 37;
Matches 9; Conservative 1; Mismatches 1; Indels 0; Gaps 0;

OY      1 CCSSSSSSSS 11
      |||||
Db      419 CCASSSSSSNS 429
      |||||

RESULT 4
G96711
unknown protein, 92941-88668 [imported] - Arabidopsis thaliana
C;Species: Arabidopsis thaliana (mouse-ear cress)
C;Date: 02-Mar-2001 #sequence_revision 02-Mar-2001 #text_change 09-Jul-2004
C;Accession: G96711
R;Theologis, A.; Ecker, J.R.; Palm, C.J.; Federspiel, N.A.; Kaul, S.; White, O.; Alonso,
Chan, C.W.; Chung, M.K.; Conn, L.; Conway, A.B.; Conway, A.R.; Creasy, T.H.; Dewar, K.;
ansen, N.F.; Hughes, B.; Huizar, L.
Nature 408, 816-820, 2000
A;Authors: Hunter, J.L.; Jenkins, J.; Johnson-Hopson, C.; Khan, S.; Khaykin, E.; Kim, C.
C.A.; Li, J.H.; Li, Y.; Lin, X.; Liu, S.X.; Liu, Z.A.; Luros, J.S.; Maiti, R.; Marziali,
Rizzo, M.; Rooney, T.; Rowley, D.; Sakano, H.; Shinn, P.; Southwick, A.M.; Sun, H.; Tallon,
A;Authors: Salzberg, S.L.; Schwarcz, J.R.; Shinn, P.; Southwick, A.M.; Sun, H.; Tallon,
ker, M.; Wu, D.; Yu, G.; Fraser, C.M.; Venter, J.C.; Davis, R.W.
A;Title: Sequence and analysis of chromosome 1 of the plant Arabidopsis.
A;Reference number: A86141; MUID:21016719; PMID:11130712
A;Accession: G96711
A;Status: preliminary
A;Molecule type: DNA
A;Residues: 1-1307 <STO>
A;Cross-references: UNIPROT:Q9S710; GB:AE005173; NID:G6524180; PIDN:AAF15073.1; GSPDB:GN
C;Genetics:
A;Gene: F14K14.18
A;Map position: 1

Query Match      79.7%; Score 47; DB 2; Length 1307;
Best Local Similarity 80.0%; Pred. No. 63;
Matches 8; Conservative 1; Mismatches 1; Indels 0; Gaps 0;

```

```

OY      1 CCCSSSSSSS 10
      |||||
Db      54 CCSTPSSSS 63
      |||||

RESULT 5
H84725
hypothetical protein At2g31850 [imported] - Arabidopsis thaliana
C;Species: Arabidopsis thaliana (mouse-ear cress)
C;Date: 02-Feb-2001 #sequence_revision 02-Feb-2001 #text_change 09-Jul-2004
C;Accession: H84725
R;Lin, X.; Kaul, S.; Roundley, S.D.; Shea, T.P.; Benito, M.I.; Town, C.D.; Fujii, C.Y.; N.
M.; Koo, H.; Moffat, K.S.; Cronin, L.A.; Shen, M.; VanAken, S.E.; Umayam, L.; Tallon, L.
eas, D.; Nierman, W.C.; White, O.; Eisen, J.A.; Salzberg, S.L.; Fraser, C.M.; Venter, J.
Nature 402, 761-768, 1999
A;Title: Sequence and analysis of chromosome 2 of the plant Arabidopsis thaliana.
A;Reference number: A84420; MUID:20083487; PMID:10617197
A;Accession: H84725
A;Status: preliminary
A;Molecule type: DNA
A;Residues: 1-113 <STO>
A;Cross-references: UNIPROT:Q9SKB5; GB:AE002093; NID:G4897751; PIDN:AAD32287.1; GSPDB:GN
C;Genetics:
A;Gene: At2g31850
A;Map position: 2

Query Match      76.3%; Score 45; DB 2; Length 113;
Best Local Similarity 63.6%; Pred. No. 22;
Matches 7; Conservative 2; Mismatches 2; Indels 0; Gaps 0;

OY      1 CCCSSSSSSSS 11
      |||||
Db      101 CCSTADSGS 111
      |||||

RESULT 6
D81551
hypothetical protein CP0675 [imported] - Chlamydothrix pneumoniae (strain AR39)
C;Species: Chlamydothrix pneumoniae, Chlamydia pneumoniae
C;Date: 31-Mar-2000 #sequence_revision 31-Mar-2000 #text_change 09-Jul-2004
C;Accession: D81551
R;Read, T.D.; Brunham, R.C.; Shen, C.; Gill, S.R.; Heidelberg, J.F.; White, O.; Hickey,
C.; Dodson, R.; Gwinn, M.; Nelson, W.; DeBoy, R.; Kolonay, J.; McClarty, G.; Salzberg,
Nucleic Acids Res. 28, 1397-1406, 2000
A;Title: Genome sequences of Chlamydia trachomatis MoPh and Chlamydia pneumoniae AR39.
A;Reference number: AB1500; MUID:20150255; PMID:10684935
A;Accession: D81551
A;Status: preliminary
A;Molecule type: DNA
A;Residues: 1-99 <REA>
A;Cross-references: UNIPROT:Q9K218; GB:AE002225; GB:AE002161; NID:G7189583; PIDN:AAF38486
C;Experimental source: strain AR39, HL cells
C;Genetics:
A;Gene: CP0675

Query Match      74.6%; Score 44; DB 2; Length 99;
Best Local Similarity 63.6%; Pred. No. 27;
Matches 7; Conservative 1; Mismatches 3; Indels 0; Gaps 0;

OY      1 CCCSSSSSSSS 11
      |||||
Db      19 CCCSDSSSEKA 29
      |||||

RESULT 7
T43486
hypothetical protein DKFP2p434N1272.1 - human (fragment)
C;Species: Homo sapiens (man)
C;Date: 21-Jan-2000 #sequence_revision 21-Jan-2000 #text_change 09-Jul-2004
C;Accession: T43486
R;Poustka, A.; Klein, M.; Mewes, H.W.; Gassenhuber, J.; Wiemann, S.
submitted to the Protein Sequence Database, December 1999

```

A;Reference number: Z22516
 A;Accession: T43486
 A;Status: preliminary
 A;Molecule type: mRNA
 A;Residues: 1-268 <AAA>
 A;Cross-references: UNIPROT:Q9P2P6; EMBL:AL133579
 A;Experimental source: adult testis; clone DKFZp434N1272
 C;Genetics:
 A;Note: DKFZp434N1272.1

Query Match 74.6%; Score 44; DB 2; Length 268;
 Best Local Similarity 80.0%; Pred. No. 51;
 Matches 8; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

QY 1 CCCSSSSSSSS 10
 |||||
 DB 34 CCCSPSSLSS 43

RESULT 8

T34329
 hypothetical protein K06A1.3 - Caenorhabditis elegans

C;Species: Caenorhabditis elegans
 C;Date: 29-Oct-1999 #sequence_revision 29-Oct-1999 #text_change 09-Jul-2004
 C;Accession: T34329
 R;Fulton, L.

submitted to the EMBL Data Library, July 1995

A;Description: The sequence of C. elegans cosmid K06A1.

A;Reference number: Z21507

A;Accession: T34329

A;Status: preliminary; translated from GB/EMBL/DBJ

A;Molecule type: DNA

A;Residues: 1-747 <FUL>

A;Cross-references: UNIPROT:Q09586; EMBL:U23449; PIDN:AAC24298.1; GSPDB:GN00020; CESP:K06A1
 A;Experimental source: strain Bristol N2; clone K06A1

C;Genetics:

A;Gene: CESP:K06A1.3

A;Map position: 2

A;Introns: 29/1; 62/3; 87/3; 133/3; 278/3; 430/3; 576/1; 659/3; 689/3

Query Match 74.6%; Score 44; DB 2; Length 747;
 Best Local Similarity 87.5%; Pred. No. 1e+02;
 Matches 7; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

QY 1 CCCSSSSS 8
 |||||
 DB 581 CCCTSSSS 588

RESULT 9

T23682

hypothetical protein M02G9.3 - Caenorhabditis elegans

C;Species: Caenorhabditis elegans

C;Date: 15-Oct-1999 #sequence_revision 15-Oct-1999 #text_change 09-Jul-2004

C;Accession: T23682

R;Matthews, L.

submitted to the EMBL Data Library, November 1996

A;Reference number: Z19781

A;Accession: T23682

A;Status: preliminary; translated from GB/EMBL/DBJ

A;Molecule type: DNA

A;Residues: 1-294 <WIL>

A;Cross-references: UNIPROT:Q9XUS0; EMBL:Z81573; PIDN:CAB04626.1; GSPDB:GN00020; CESP:M02G9

A;Experimental source: clone M02G9

C;Genetics:

A;Gene: CESP:M02G9.3

A;Map position: 2

A;Introns: 20/3; 76/2; 182/3; 223/3

Query Match 72.9%; Score 43; DB 2; Length 294;
 Best Local Similarity 88.9%; Pred. No. 72;
 Matches 8; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

QY 2 CCSSSSSSS 10
 |||||
 DB 42 CCSSSSSSS 50

RESULT 10

S03631
 homeotic protein Scr - fruit fly (Drosophila melanogaster)

N;Alternate names: sex combs reduced protein

C;Species: Drosophila melanogaster

C;Date: 10-Sep-1999 #sequence_revision 10-Sep-1999 #text_change 09-Jul-2004

C;Accession: S03631

R;LeMotte, P.K.; Kuroiwa, A.; Fessler, L.I.; Gehring, W.J.

EMBO J. 8, 219-227, 1989

A;Title: The homeotic gene sex combs reduced of Drosophila: gene structure and embryonic

A;Reference number: S03631; MUID:89231621; PMID:2565809

A;Accession: S03631

A;Molecule type: DNA

A;Residues: 1-413 <LEM>

A;Cross-references: UNIPROT:P09077; EMBL:X14475

C;Genetics:

A;Gene: Scr

A;Cross-references: FlyBase:FBgn000339

C;Superfamily: homeotic protein Scr; homeobox homology

C;Keywords: DNA binding; homeobox; nucleus; transcription regulation

F;321-377/Domain: homeobox homology <HOX>

Query Match 72.9%; Score 43; DB 1; Length 413;
 Best Local Similarity 72.7%; Pred. No. 90;
 Matches 8; Conservative 2; Mismatches 1; Indels 0; Gaps 0;

QY 1 CCCSSSSSSS 11
 |||||
 DB 119 CSCSSSSSSNS 129

RESULT 11

S47300

gene F protein - rinderpest virus

C;Species: rinderpest virus

C;Date: 06-Jan-1995 #sequence_revision 06-Jan-1995 #text_change 09-Jul-2004

C;Accession: S47300; PQ0865

R;Evans, S.A.; Baron, M.D.; Chamberlain, R.W.; Goatley, L.; Barrett, T.

submitted to the EMBL Data Library, March 1994

A;Description: The complete nucleotide sequence of the fusion protein gene of the vaccine

A;Reference number: S47299

A;Accession: S47300

A;Molecule type: DNA

A;Residues: 1-546 <EVA>

A;Cross-references: UNIPROT:P41360; EMBL:Z31656; NID:G535406; PIDN:CAA83482.1; PID:G53540

R;Chamberlain, R.W.; Wamwayi, H.M.; Hockley, E.; Shaila, M.S.; Goatley, L.; Knowles, N.J.

J. Gen. Virol. 74, 2775-2780, 1993

A;Title: Evidence for different lineages of rinderpest virus reflecting their geographic

A;Reference number: PQ0865; MUID:94103786; PMID:8277286

A;Accession: PQ0865

A;Molecule type: mRNA

A;Residues: 86-191 <CHA>

C;Genetics:

A;Gene: F

C;Superfamily: parainfluenza virus cell fusion protein

C;Keywords: glycoprotein; membrane fusion; transmembrane protein

Query Match 72.9%; Score 43; DB 2; Length 546;

Best Local Similarity 70.0%; Pred. No. 1.1e+02;

Matches 7; Conservative 0; Mismatches 3; Indels 0; Gaps 0;

QY 1 CCCSSSSSSS 10
 |||||
 DB 511 CCCKSSSDS 520

RESULT 12

I38228

Shb protein - human

C;Species: Homo sapiens (man)
C;Date: 29-May-1998 #sequence_revision 29-May-1998 #text_change 09-Jul-2004
C;Accession: I38228; S37577
R;Welsh, M.; Mares, J.; Karlsson, T.; Lavergne, C.; Breant, B.; Claesson-Welsh, L.
Oncogene 9, 19-27, 1994
A;Title: Shb is a ubiquitously expressed Src homology 2 protein.
A;Reference number: I38228; MUID:941134414; PMID:8302579

C;Accession: I38228
A;Status: preliminary; translated from GB/EMBL/DBJ

A;Molecule type: mRNA
A;Residues: 1-596 <RES>

A;Cross-references: UNIPROT:Q15464; EMBL:X75342; NID:g406737; PIDN:CAA53091.1; PID:g4067

R;Welsh, M.

submitted to the EMBL Data Library, October 1993

A;Reference number: S37577

A;Accession: S37577

A;Status: preliminary

A;Molecule type: mRNA

A;Residues: 1-596 <WEL>

C;Superfamily: Shb protein; SH2 homology

F;497-595/Domain: SH2 homology <SH21>

Query Match 72.9%; Score 43; DB 2; Length 596;

Best Local Similarity 58.8%; Pred. No. 1.1e+02;

Matches 10; Conservative 1; Mismatches 0; Indels 6; Gaps 1;

Qy 1 CCCSSS-----SSSS 11

Db 226 CCCASSGAGAAASSSS 242

RESULT 13

T00282

hypothetical protein - Escherichia coli plasmid p0157

C;Species: Escherichia coli

C;Date: 01-Feb-1999 #sequence_revision 01-Feb-1999 #text_change 09-Jul-2004

C;Accession: T00282

R;Makino, K.; Ishii, K.; Yaeunaga, T.; Hattori, M.; Yokoyama, K.; Yatsudo, H.C.; Kubota,

S.; Shinagawa, H.

DNA Res. 5, 1-9, 1998

A;Title: Complete nucleotide sequences of 93-kb and 3.3-kb plasmids of an enterohemorrhagic

A;Reference number: Z14127; MUID:98290540; PMID:9628576

A;Accession: T00282

A;Status: preliminary; translated from GB/EMBL/DBJ

A;Molecule type: DNA

A;Residues: 1-129 <MAK>

A;Cross-references: UNIPROT:O82903; EMBL:AB011549; NID:g4589740; PIDN:BAA31801.1; PID:g3

A;Experimental source: strain EHEC O157:H7, substrain RIMD 0509952

C;Genetics:

A;Genome: plasmid p0157

Query Match 71.2%; Score 42; DB 2; Length 129;

Best Local Similarity 81.8%; Pred. No. 55;

Matches 9; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

Qy 1 CCCSSSSSSSS 11

Db 18 CYCSSSFSSSS 28

RESULT 14

S18946

ultra high-sulfur keratin 1 - human

N;Alternate names: UHS Keratin; ultra high-sulfur matrix protein

C;Species: Homo sapiens (man)

C;Date: 22-Oct-1999 #sequence_revision 22-Oct-1999 #text_change 09-Jul-2004

C;Accession: S18946; B36686

R;Drabant, B.; Doenecke, D.

submitted to the EMBL Data Library, December 1991

A;Description: Nucleotide sequence of a Human high-sulphur keratin cDNA.

A;Reference number: S18946

A;Accession: S18946

A;Molecule type: mRNA

A;Residues: 1-169 <DRA>

A;Cross-references: UNIPROT:Q14564; EMBL:X63755; NID:g32471; PIDN:CAA45283.1; PID:g32472

R;Wackinnon, P.J.; Powell, B.C.; Rogers, G.E.

J. Cell Biol. 111, 2587-2600, 1990

A;Title: Structure and expression of genes for a class of cysteine-rich proteins of the

A;Reference number: A36686; MUID:91115951; PMID:1703541

A;Accession: B36686

A;Molecule type: DNA

A;Residues: 1-39, 'Y', 41-169 <MAC>

A;Cross-references: GB:X55293; NID:g34078; PIDN:CAA39005.1; PID:g34079

C;Genetics:

A;Gene: GDB:KRN1

A;Cross-references: GDB:125257; OMIM:148021

A;Map position: 11q13-11q13

C;Superfamily: ultra-high-sulfur keratin

C;Keywords: hair; tandem repeat

F;7-15/Region: Ser-rich nonapeptide repeat

F;59-68/Region: Gly-rich decapeptide repeat

F;69-78/Region: Gly-rich decapeptide repeat

F;79-88/Region: Cys-rich decapeptide repeat

F;89-97/Region: Ser-rich nonapeptide repeat

F;98-107/Region: Cys-rich decapeptide repeat

F;108-117/Region: Cys-rich decapeptide repeat

F;118-126/Region: Ser-rich nonapeptide repeat

F;127-136/Region: Cys-rich decapeptide repeat

F;137-145/Region: Ser-rich nonapeptide repeat

F;146-155/Region: Cys-rich decapeptide repeat

F;156-165/Region: Cys-rich decapeptide repeat

Query Match 71.2%; Score 42; DB 1; Length 169;

Best Local Similarity 70.0%; Pred. No. 66;

Matches 7; Conservative 0; Mismatches 3; Indels 0; Gaps 0;

Qy 1 CCCSSSSSSS 10

Db 86 CCCSSGCGSS 95

RESULT 15

T41987

hypothetical protein U86 - human herpesvirus 7 (strain J1)

C;Species: human herpesvirus 7

A;Variety: strain J1

C;Date: 03-Dec-1999 #sequence_revision 03-Dec-1999 #text_change 09-Jul-2004

C;Accession: T41987

R;Nicholas, J.

submitted to the EMBL Data Library, December 1995

A;Description: Determination and analysis of the complete nucleotide sequence of human he

A;Reference number: Z22022

A;Accession: T41987

A;Status: preliminary; translated from GB/EMBL/DBJ

A;Molecule type: DNA

A;Residues: 1-1205 <NIC>

A;Cross-references: UNIPROT:Q69513; EMBL:U43400; PIDN:AAC54747.1

A;Experimental source: strain J1

C;Genetics:

A;Note: U86

Query Match 71.2%; Score 42; DB 2; Length 1205;

Best Local Similarity 72.7%; Pred. No. 2.4e+02;

Matches 8; Conservative 1; Mismatches 2; Indels 0; Gaps 0;

Qy 1 CCCSSSSSSSS 11

Db 677 CSCSSSSSASS 687

Search completed: November 15, 2005, 12:23:31

Job time : 10.8421 secs

GenCore version 5.1.6
Copyright (c) 1993 - 2005 Compugen Ltd.

OM protein - protein search, using sw model

Run on: November 15, 2005, 12:12:49 ; Search time 80.1842 Seconds
(without alignments)
70.249 Million cell updates/sec

Title: US-10-820-998-4
Perfect score: 59
Sequence: 1 SSSSSSSSSCCC 11

Scoring table: BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 1612378 seqs, 512079187 residues

Total number of hits satisfying chosen parameters: 1612378

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database : UniProt_03.*

1: uniprot_sprot.*

2: uniprot_trembl.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	59	100.0	47	2	Q76NT4 dictyosteli
2	54	91.5	4261	2	Q81FP4 plasmodium
3	50	84.7	240	2	Q6ZAY7 oryza sativ
4	50	84.7	351	2	Q6VTN6 drosophila
5	50	84.7	360	2	Q9SD60 arabidopsis
6	50	84.7	504	2	Q9AXJ9 zinnia eleg
7	50	84.7	505	2	Q9AXJ8 zinnia eleg
8	50	84.7	974	2	P93155 gossypium h
9	50	84.7	974	2	Q8WIW0 gossypium h
10	50	84.7	978	2	Q9AXK0 zinnia eleg
11	49	83.1	922	2	Q8SZ15 drosophila
12	49	83.1	155	2	Q80WU5 mus musculus
13	48.5	82.2	436	2	Q6PFM7 brachydanio
14	47	79.7	240	2	Q9LKK0 atriplex ho
15	47	79.7	367	2	Q6AV38 oryza sativ
16	47	79.7	630	2	Q9XHQ4 hordeum vul
17	47	79.7	958	2	Q9SN37 arabidopsis
18	47	79.7	985	2	Q8LPK5 arabidopsis
19	47	79.7	985	2	Q9CSZ8 arabidopsis
20	47	79.7	1670	2	Q23901 dictyosteli
21	47	79.7	1671	2	Q869S5 dictyosteli
22	46	78.0	198	2	Q95WU4 giardia lam
23	46	78.0	198	2	Q7RI55 giardia lam
24	46	78.0	198	2	Q9BI02 giardia lam
25	46	78.0	216	1	P15744 eimeria ten
26	46	78.0	313	2	Q9XWR2 caenorhabdi
27	45.5	77.1	440	2	Q90X48 brachydanio
28	45.5	77.1	518	2	Q640H3 xenopus lae
29	45	76.3	1205	2	O56300 human herpe
30	45	76.3	1205	2	Q69513 human herpe
31	44	74.6	99	1	NIC1_HUMAN

32 44 74.6 131 1 YJG4 YEAST
33 44 74.6 328 2 Q6ZNM3
34 44 74.6 405 2 Q7XGR1
35 44 74.6 649 2 Q8CAD4
36 44 74.6 738 2 Q8RT92
37 44 74.6 841 2 Q8CCH4
38 44 74.6 943 2 Q8IBA9
39 44 74.6 1923 2 Q8IBH8
40 44 74.6 5415 2 Q8IE74
41 43 72.9 162 2 Q8BWC8
42 43 72.9 185 2 O65658
43 43 72.9 193 1 TSA1 MOUSE
44 43 72.9 208 2 Q6LJ59
45 43 72.9 217 2 Q9SCY3

ALIGNMENTS

RESULT 1
Q76NT4 PRELIMINARY; PRT; 47 AA.
ID O76NT4
AC O76NT4;
DT 05-JUL-2004 (TrEMBLrel. 27, Created)
DT 05-JUL-2004 (TrEMBLrel. 27, Last sequence update)
DT 05-JUL-2004 (TrEMBLrel. 27, Last annotation update)
DE Hypothetical protein.
OS Dictyostelium discoideum (Slime mold).
OC Eukaryota; Mycetozoa; Dictyostelida; Dictyostelium.
OX NCBI_TaxID=44689;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=AX4;
RX MEDLINE=22092622; PubMed=12097910; DOI=10.1038/nature00847;
RA Gloeckner G., Eichinger L., Szafranski K., Pachebat J., Dear P.,
RA Lehmann R., Baumgart C., Parra G., April J.F., Guigo R., Kumpf K.,
RA Tunggal B., Cox E., Quail M.A., Platzer M., Rosenthal A., Noegel A.A.;
RT "Sequence and analysis of chromosome 2 of Dictyostelium discoideum.";
RL Nature 418:79-85(2002).
RN [2]
RP SEQUENCE FROM N.A.
RC STRAIN=AX4;
RX Baumgart C.;

RL Submitted (FEB-2004) to the EMBL/GenBank/DBJ databases.

DR EMBL; AC117081; AAS38828.1; -

KW Hypothetical protein.

SQ SEQUENCE 47 AA; 4489 MW; 3781AB1651807149 CRC64;

Query Match 100.0%; Score 59; DB 2; Length 47;

Best Local Similarity 100.0%; Pred. No. 0.29;

Matches 11; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 SSSSSSSSSCCC 11

DB 30 SSSSSSSSSCCC 40

RESULT 2

Q81FP4

ID Q81FP4 PRELIMINARY; PRT; 4261 AA.

AC Q81FP4;

DT 01-MAR-2003 (TrEMBLrel. 23, Created)

DT 01-MAR-2003 (TrEMBLrel. 23, Last sequence update)

DT 01-OCT-2003 (TrEMBLrel. 25, Last annotation update)

DE Erythrocyte membrane-associated antigen, putative.

GN Name=PF01045c;

OS Plasmodium falciparum (Isolate 3D7).

OC Eukaryota; Alveolata; Apicomplexa; Haemosporida; Plasmodium.

OX NCBI_TaxID=36329;

RN [1]

RP SEQUENCE FROM N.A.

RA Hamlin N., Pain A., Berriman B., Hall N., Bowman S., Churcher C.,

RA Harris B., Harris D., Lawson D., Quail M., Barrell B.;

```

RL Submitted (FEB-1999) to the EMBL/GenBank/DBJ databases.
DR EMBL; AL035477; CAD49256.1; -
DR GO; GO:0003676; F:nucleic acid binding; IEA.
DR InterPro; IPR001201; PAP_25A_core.
SQ SEQUENCE 4261 AA; 508568 MW; 4F9F041BC21E44D8 CRC64;

Query Match 91.5%; Score 54; DB 2; Length 4261;
Best Local Similarity 90.9%; Pred. No. 89;
Matches 10; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

Qy 1 SSSSSSSSSCCC 11
Db 3433 SCSSSSSSCCC 3443

RESULT 3
ID Q6ZAY7 PRELIMINARY; PRT; 240 AA.
AC Q6ZAY7;
DT 05-JUL-2004 (TrEMBLrel. 27, Created)
DT 05-JUL-2004 (TrEMBLrel. 27, Last sequence update)
DT 05-JUL-2004 (TrEMBLrel. 27, Last annotation update)
DE Hypothetical protein OUL540_G08.6.
GN Name=OUL540_G08.6;
OS Oryza sativa (japonica cultivar-group).
OC Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
OC Spermatophyta; Magnoliophyta; Liliopsida; Poales; Poaceae;
OC Ehrhartoideae; Oryzoae; Oryza.
OX NCBI_TaxID=39947;
RN [1]
RP SEQUENCE FROM N.A.
RA Sasaki T., Matsumoto T., Yamamoto K.;
RL Submitted (JAN-2002) to the EMBL/GenBank/DBJ databases.
DR EMBL; AP004650; BAC99398.1; -
KW Hypothetical protein.
SQ SEQUENCE 240 AA; 26540 MW; 7284738FA9BAE445 CRC64;

Query Match 84.7%; Score 50; DB 2; Length 240;
Best Local Similarity 81.8%; Pred. No. 19;
Matches 9; Conservative 1; Mismatches 1; Indels 0; Gaps 0;

Qy 1 SSSSSSSSSCCC 11
Db 17 SFSSASSSSCCC 27

RESULT 4
ID Q6VYN6 PRELIMINARY; PRT; 351 AA.
AC Q6VYN6;
DT 05-JUL-2004 (TrEMBLrel. 27, Created)
DT 05-JUL-2004 (TrEMBLrel. 27, Last sequence update)
DT 05-JUL-2004 (TrEMBLrel. 27, Last annotation update)
DE Scute.
OS Drosophila virilis (Fruit fly).
OC Eukaryota; Metazoa; Arthropoda; Hexapoda; Insecta; Pterygota;
OC Neoptera; Endopterygota; Diptera; Brachycera; Muscomorpha;
OC Ephydroidea; Drosophilidae; Drosophila.
OX NCBI_TaxID=7244;
RN [1]
RP SEQUENCE FROM N.A.
RA Wischnik L.A., Timmer J.R., Megna L.A., Cline T.W.;
RX PubMed=14704182;
RT "Recruitment of the proneural gene scute to the Drosophila sex-
RT determination pathway.";
RL Genetics 165:2007-2027(2003).
CC -!- SIMILARITY: Contains 1 basic helix-loop-helix (bHLH) domain.
DR EMBL; AY319375; AA082728.1; -
DR InterPro; IPR001092; HLH_basic.
DR Pfam; PF00010; HLH; 1.
DR SMART; SM00353; HLH; 1.
DR PROSITE; PS00888; HLH; 1.
SQ SEQUENCE 351 AA; 37567 MW; A08C2145ADE26460 CRC64;

Query Match 84.7%; Score 50; DB 2; Length 351;
Best Local Similarity 100.0%; Pred. No. 28;
Matches 10; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 SSSSSSSSSCCC 10
Db 75 SSSSSSSSSCCC 84

RESULT 5
ID Q9SD60 PRELIMINARY; PRT; 360 AA.
AC Q9SD60;
DT 01-MAY-2000 (TrEMBLrel. 13, Created)
DT 01-MAY-2000 (TrEMBLrel. 13, Last sequence update)
DT 01-OCT-2003 (TrEMBLrel. 25, Last annotation update)
DE Hypothetical protein FI3112.180.
GN Name=FI3112.180;
OS Arabidopsis thaliana (Mouse-ear cress).
OC Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
OC Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots; rosids;
OC eurosids II; Brassicales; Brassicaceae; Arabidopsids.
OX NCBI_TaxID=3702;
RN [1]
RP SEQUENCE FROM N.A.
RA Choisne N., Robert C., Brottier P., Wincker P., Cattolico L.,
RA Artiguenave F., Saurin W., Weissenbach J., Mewes H.W., Lemcke K.,
RA Mayer K.F.X., Quetier F., Salanoubat M.;
RL Submitted (NOV-1999) to the EMBL/GenBank/DBJ databases.
RN [2]
RP SEQUENCE FROM N.A.
RA EU Arabidopsis sequencing project;
RL Submitted (JAN-1999) to the EMBL/GenBank/DBJ databases.
DR EMBL; ALU13292; CAB61959.1; -
DR PIR; T45649; T45649.
DR InterPro; IPR006527; F_box_assoc_1.
DR TIGRPFAMS; TIGR01640; F_box_assoc_1; 1.
KW Hypothetical protein.
SQ SEQUENCE 360 AA; 41153 MW; A7F21DBB44EB64D6 CRC64;

Query Match 84.7%; Score 50; DB 2; Length 360;
Best Local Similarity 100.0%; Pred. No. 28;
Matches 10; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 SSSSSSSSSCCC 10
Db 97 SSSSSSSSSCCC 106

RESULT 6
ID Q9AXJ9 PRELIMINARY; PRT; 504 AA.
AC Q9AXJ9;
DT 01-JUN-2001 (TrEMBLrel. 17, Created)
DT 01-JUN-2001 (TrEMBLrel. 17, Last sequence update)
DT 01-OCT-2003 (TrEMBLrel. 25, Last annotation update)
DE Cellulose synthase Cesa-2 (Fragment).
OS Zinnia elegans (Zinnia).
OC Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
OC Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots; asterids;
OC campanulids; Asterales; Asteraceae; Heliantheae; Zinnia.
OX NCBI_TaxID=34245;
RN [1]
RP SEQUENCE FROM N.A.
RX MEDLINE=21438255; PubMed=11554477; DOI=10.1023/A:1010615027986;
RA Haigler C.H., Ivanova-Datcheva M., Hogan P.S., Salnikov V.V.,
RA Hwang S., Martin K., Delmer D.P.;
RT "Carbon partitioning to cellulose synthesis.";
RL Plant Mol. Biol. 47:29-51(2001).
DR EMBL; AF323040; AAK1589.1; -
DR GO; GO:0016020; C:membrane; IEA.
DR GO; GO:0016760; F:cellulose synthase (UDP-forming) activity; IEA.

```

DR GO:0030244; P:cellulose biosynthesis; IEA.
 DR InterPro; IPR005150; Cellulose synth.
 DR Pfam; PF03552; Cellulose_synt; 1.
 FT NON TER 1
 SQ SEQUENCE 504 AA; 56578 MW; D5C3BE2E1565CA1D CRC64;
 Query Match 84.7%; Score 50; DB 2; Length 504;
 Best Local Similarity 90.0%; Pred. No. 39;
 Matches 9; Conservative 0; Mismatches 1; Indels 0; Gaps 0;
 QY 2 SSSSSSSCCC 11
 Db 83 SPSSSSSSCCC 92
 RESULT 7
 Q9AXJ8 PRELIMINARY; PRT; 505 AA.
 AC Q9AXJ8;
 DT 01-JUN-2001 (TrEMBLrel. 17, Created)
 DT 01-JUN-2001 (TrEMBLrel. 17, Last sequence update)
 DT 01-OCT-2003 (TrEMBLrel. 25, Last annotation update)
 DE Cellulose synthase Cesa-3 (fragment).
 OS Zinnia elegans (Zinnia).
 OC Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
 OC Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots; asterids;
 OC campanulids; Asterales; Asteraceae; Malvaceae; Malvoideae; Heliantheae; Zinnia.
 OX NCBI_TaxID=34245;
 RN [1]
 SEQUENCE FROM N.A.
 RP MEDLINE=21438255; PubMed=11554477; DOI=10.1023/A:1010615027986;
 RA Haigler C.H., Ivanova-Datcheva M., Hogan P.S., Salnikov V.V.,
 RA Huang S., Martin K., Delmer D.P.;
 RT "Carbon partitioning to cellulose synthesis.";
 RL Plant Mol. Biol. 47:29-51(2001).
 DR EMBL; AF323041; AAK11590.1; -;
 DR GO:0016020; C:membrane; IEA.
 DR GO:0016760; F:cellulose synthase (UDP-forming) activity; IEA.
 DR GO:0030244; P:cellulose biosynthesis; IEA.
 DR InterPro; IPR005150; Cellulose synth.
 DR Pfam; PF03552; Cellulose_synt; 1.
 FT NON TER 1
 SQ SEQUENCE 505 AA; 56478 MW; 9400BA486AD1B02D CRC64;
 Query Match 84.7%; Score 50; DB 2; Length 505;
 Best Local Similarity 90.0%; Pred. No. 39;
 Matches 9; Conservative 0; Mismatches 1; Indels 0; Gaps 0;
 QY 2 SSSSSSSCCC 11
 Db 83 SPSSSSSSCCC 92
 RESULT 8
 P93155 PRELIMINARY; PRT; 974 AA.
 AC P93155;
 DT 01-MAY-1997 (TrEMBLrel. 03, Created)
 DT 01-MAY-1997 (TrEMBLrel. 03, Last sequence update)
 DT 25-OCT-2004 (TrEMBLrel. 28, Last annotation update)
 DE Cellulose synthase.
 GN Names=celA1; Synonyms=CesA1;
 OS Gossypium hirsutum (Upland cotton).
 OC Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
 OC Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots; rosids;
 OC eurosids II; Malvales; Malvaceae; Malvoideae; Gossypium.
 OX NCBI_TaxID=3635;
 RN [1]
 SEQUENCE FROM N.A.
 RP STRAIN=Acala SJ-2; TISSUE=Fiber;
 RC MEDLINE=97057296; PubMed=8901633; DOI=10.1073/pnas.93.22.12637;
 RX Pear J.R., Kawagoe Y., Schreckengost W.E., Delmer D.P., Stalker D.M.;
 RA "Higher plants contain homologs of the bacterial celA genes encoding

RT the catalytic subunit of cellulose synthase.";
 RL Proc. Natl. Acad. Sci. U.S.A. 93:12637-12642(1996).
 RN [2]
 SEQUENCE FROM N.A.
 RP PubMed=15256507; DOI=10.1101/gr.2673204;
 RA Grover C.E., Kim H., Wing R.A., Paterson A.H., Wendel J.F.;
 RT "Incongruent Patterns of Local and Global Genome Size Evolution in Cotton";
 RL Genome Res. 14:1474-1482(2004).
 DR EMBL; U58283; AAB37766.1; -;
 DR EMBL; AY632360; AAT64028.1; -;
 DR PIR; T10797; T10797.
 DR GO:0016020; C:membrane; IEA.
 DR GO:0016760; F:cellulose synthase (UDP-forming) activity; IEA.
 DR GO:0030244; P:cellulose biosynthesis; IEA.
 DR InterPro; IPR005150; Cellulose synth.
 DR Pfam; PF03552; Cellulose_synt; 1.
 SQ SEQUENCE 974 AA; 109703 MW; 2656CA2D05D100A7 CRC64;
 Query Match 84.7%; Score 50; DB 2; Length 974;
 Best Local Similarity 90.0%; Pred. No. 73;
 Matches 9; Conservative 0; Mismatches 1; Indels 0; Gaps 0;
 QY 2 SSSSSSSCCC 11
 Db 553 SSSSSSSCCC 562
 RESULT 9
 Q8W1W0 PRELIMINARY; PRT; 974 AA.
 AC Q8W1W0;
 DT 01-MAR-2002 (TrEMBLrel. 20, Created)
 DT 01-MAR-2002 (TrEMBLrel. 20, Last sequence update)
 DT 01-OCT-2003 (TrEMBLrel. 25, Last annotation update)
 DE Cellulose synthase A4.
 GN Name=Cesa4;
 OS Gossypium hirsutum (Upland cotton).
 OC Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
 OC Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots; rosids;
 OC eurosids II; Malvales; Malvaceae; Malvoideae; Gossypium.
 OX NCBI_TaxID=3635;
 RN [1]
 SEQUENCE FROM N.A.
 RP MEDLINE=21608363; PubMed=11743074; DOI=10.1104/pp.127.4.1361;
 RX Kim H.J., Triplett B.A.;
 RT "Cotton fiber growth in planta and in vitro. Models for plant cell elongation and cell wall biogenesis.";
 RL Plant Physiol. 127:1361-1366(2001).
 DR EMBL; AF413210; AAL37718.1; -;
 DR GO:0016020; C:membrane; IEA.
 DR GO:0016760; F:cellulose synthase (UDP-forming) activity; IEA.
 DR GO:0030244; P:cellulose biosynthesis; IEA.
 DR Pfam; PF03552; Cellulose_synt; 1.
 SQ SEQUENCE 974 AA; 109499 MW; F8AFB66ABA2B8071 CRC64;
 Query Match 84.7%; Score 50; DB 2; Length 974;
 Best Local Similarity 90.0%; Pred. No. 73;
 Matches 9; Conservative 0; Mismatches 1; Indels 0; Gaps 0;
 QY 2 SSSSSSSCCC 11
 Db 553 SSSSSSSCCC 562
 RESULT 10
 Q9AXK0 PRELIMINARY; PRT; 978 AA.
 AC Q9AXK0;
 DT 01-JUN-2001 (TrEMBLrel. 17, Created)
 DT 01-JUN-2002 (TrEMBLrel. 21, Last sequence update)
 DT 01-OCT-2003 (TrEMBLrel. 25, Last annotation update)
 DE Cellulose synthase Cesa-1.

```

OS Zinnia elegans (Zinnia).
OC Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
OC Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots; asterids;
OC Campanulids; Asterales; Asteraceae; Asteroideae; Heliantheae; Zinnia.
OX NCBI_TaxID=34245;
RN [1]
RN SEQUENCE FROM N.A.
RX MEDLINE=21438255; PubMed=11554477; DOI=10.1023/A:1010615027986;
RA Haigler C.H., Ivanova-Datcheva M., Hogan P.S., Salnikov V.V.,
RA Hwang S., Martin K., Delmer D.P.;
RT "Carbon partitioning to cellulose synthesis.";
RL Plant Mol. Biol. 47:29-51(2001).
RN [2]
RN SEQUENCE FROM N.A.
RA Hwang S., Haigler C.H.;
RL Submitted (FEB-2002) to the EMBL/GenBank/DBJ databases.
DR EMBL; AF323039; AAK11588.2; -
DR GO; GO:0016020; C:membrane; IEA.
DR GO; GO:0016760; F:cellulose synthase (UDP-forming) activity; IEA.
DR GO; GO:0030244; F:cellulose biosynthesis; IEA.
DR InterPro; IPR005150; Cellulose synth.
DR Pfam; PF03552; Cellulose synt.; 1.
SQ SEQUENCE 378 AA; 110053 MW; 399D76EF3427B1D CRC64;

Query Match 84.7%; Score 50; DB 2; Length 978;
Best Local Similarity 90.0%; Pred. No. 73;
Matches 9; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

Qy 2 SSSSSSSCCC 11
Db 557 SPSSSSSSCC 566

RESULT 11
Q8SZ15 PRELIMINARY; PRT; 92 AA.
AC Q8SZ15;
DT 01-JUN-2002 (TrEMBLrel. 21, Created)
DT 01-JUN-2002 (TrEMBLrel. 21, Last sequence update)
DT 01-OCT-2002 (TrEMBLrel. 22, Last annotation update)
DE L006106p.
GN ORFNames=CG13847;
OS Drosophila melanogaster (Fruit fly).
OC Eukaryota; Metazoa; Arthropoda; Hexapoda; Insecta; Pterygota;
OC Neoptera; Endopterygota; Diptera; Brachycera; Muscomorpha;
OC Ephydroidea; Drosophilidae; Drosophila.
OX NCBI_TaxID=7227;
RN [1]
RN SEQUENCE FROM N.A.
RC STRAIN=Berkley;
RA Stapleton M., Brokstein P., Hong L., Agbavani A., Carlson J.,
RA Champe M., Chavez C., Dorsett V., Dresnek D., Farfan D., Frise E.,
RA George R., Gonzalez M., Guarin H., Krommiller B., Li P., Liao G.,
RA Miranda A., Mungall C.J., Nunoo J., Pacleb J., Paragas V., Park S.,
RA Patel S., Phouanavong S., Wan K., Yu C., Lewis S.E., Rubin G.M.,
RA Celnikier S.;
RL Submitted (DEC-2001) to the EMBL/GenBank/DBJ databases.
DR EMBL; AY070873; AAL48495.1; -
DR FlyBase; FBgn038967; CG13847.
SQ SEQUENCE 92 AA; 10308 MW; CSA24EABC43137E5 CRC64;

Query Match 83.1%; Score 49; DB 2; Length 92;
Best Local Similarity 80.0%; Pred. No. 11;
Matches 8; Conservative 2; Mismatches 0; Indels 0; Gaps 0;

Qy 2 SSSSSSSCCC 11
Db 51 NSSSSSSCCC 60

RESULT 12
Q80WU5 PRELIMINARY; PRT; 155 AA.
ID Q80WU5

```

```

AC Q80WU5;
DT 01-JUN-2003 (TrEMBLrel. 24, Created)
DT 01-JUN-2003 (TrEMBLrel. 24, Last sequence update)
DT 01-JUN-2003 (TrEMBLrel. 24, Last annotation update)
DE Similar to dachshund 1 (Drosophila).
OS Mus musculus (Mouse)
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
OX NCBI_TaxID=10090;
RN [1]
RN SEQUENCE FROM N.A.
RC STRAIN=C57BL/6; TISSUE=Brain;
RX MEDLINE=22388257; PubMed=12477932; DOI=10.1073/pnas.242603899;
RA Strausberg R.L., Feingold E.A., Grouse L.H., Derge J.G.,
RA Klausner R.D., Collins F.S., Wagner L., Shenmen C.M., Schuler G.D.,
RA Altschul S.F., Zeeberg B., Buetow K.H., Schaefer C.F., Bhat N.K.,
RA Hopkins R.F., Jordan H., Moore T., Max S.I., Wang J., Hsieh F.,
RA Diatchenko L., Marusina K., Farmer A.A., Rubin G.M., Hong L.,
RA Stapleton M., Soares M.B., Bonaldo M.F., Casavant T.L., Scheetz T.E.,
RA Brownstein M.J., Usdin T.B., Toshiyuki S., Carninci P., Prange C.,
RA Raha S.S., Loquellano N.A., Peters G.J., Abramson R.D., Mullaly S.J.,
RA Bosak S.A., McEwan P.J., McKernan K.J., Malek J.A., Gunaratne P.H.,
RA Richards S., Worley K.C., Hale S., Garcia A.M., Gay L.J., Hulyk S.W.,
RA Villalón D.K., Muzny D.M., Sodergren E.J., Lu X., Gibbs R.A.,
RA Fahey J., Helton E., Kettelman M., Madan A., Rodrigues S., Sanchez A.,
RA Whiting M., Madan A., Young A.C., Shevchenko Y., Bouffard G.G.,
RA Blakesley R.W., Touchman J.W., Green E.D., Dickson M.C.,
RA Rodriguez A.C., Grimwood J., Schmutz J., Myers R.M., Butterfield Y.S.,
RA Krzywinski M.I., Skaleka U., Smailus D.E., Schnerch A., Schein J.E.,
RA Jones S.J., Marra M.A.;
RT "Generation and initial analysis of more than 15,000 full-length human
and mouse cDNA sequences.";
RL Proc. Natl. Acad. Sci. U.S.A. 99:16899-16903(2002).
RN [2]
RN SEQUENCE FROM N.A.
RC STRAIN=C57BL/6; TISSUE=Brain;
RA Strausberg R.;
RL Submitted (MAY-2003) to the EMBL/GenBank/DBJ databases.
DR EMBL; BC052005; AAH52005.1; -
SQ SEQUENCE 155 AA; 13662 MW; CD1B054DBC769450 CRC64;

Query Match 83.1%; Score 49; DB 2; Length 155;
Best Local Similarity 90.9%; Pred. No. 17;
Matches 10; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

Qy 1 SSSSSSSSSCCC 11
Db 136 SSSSSSSSSCSC 146

RESULT 13
Q6PFM7 PRELIMINARY; PRT; 436 AA.
AC Q6PFM7;
DT 05-JUL-2004 (TrEMBLrel. 27, Created)
DT 05-JUL-2004 (TrEMBLrel. 27, Last sequence update)
DT 05-JUL-2004 (TrEMBLrel. 27, Last annotation update)
DE Solute carrier family 35, member A5.
GN Name=slc35a5;
OS Brachydanio rerio (Zebrafish) (Danio rerio).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Actinopterygii; Neopterygii; Teleostei; Ostariophysi; Cypriniformes;
OC Cyprinidae; Danio.
OX NCBI_TaxID=7955;
RN [1]
RN SEQUENCE FROM N.A.
RC STRAIN=AB; TISSUE=Whole body;
RX MEDLINE=22388257; PubMed=12477932; DOI=10.1073/pnas.242603899;
RA Strausberg R.L., Feingold E.A., Grouse L.H., Derge J.G.,
RA Klausner R.D., Collins F.S., Wagner L., Shenmen C.M., Schuler G.D.,
RA Altschul S.F., Zeeberg B., Buetow K.H., Schaefer C.F., Bhat N.K.,
RA Hopkins R.F., Jordan H., Moore T., Max S.I., Wang J., Hsieh F.,
RA Diatchenko L., Marusina K., Farmer A.A., Rubin G.M., Hong L.,

```

RA Stapleton M., Soares M.B., Bernaldo M.F., Casavant T.L., Scheetz T.E.,
 RA Brownstein M.J., Uedin T.B., Toshiyuki S., Carninci P., Prange C.,
 RA Raha S.S., Loquellano N.A., Peters G.J., Abramson R.D., Mullaly S.J.,
 RA Bosak S.A., McSwan P.J., McKernan K.J., Malek J.A., Gunaratne P.H.,
 RA Richards S., Worley K.C., Hale S., Garcia A.M., Gay L.J., Hulyk S.W.,
 RA Villalon D.K., Muzny D.M., Sodergren E.J., Lu X., Gibbs R.A.,
 RA Fahey J., Helton E., Kettelman M., Madan A., Rodrigues S., Sanchez A.,
 RA Whiting M., Madan A., Young A.C., Shevchenko Y., Bouffard G.G.,
 RA Blakeley R.W., Touchman J.W., Green E.D., Dickson M.C.,
 RA Rodriguez A.C., Grimwood J., Schmutz J., Myers R.M., Butterfield Y.S.,
 RA Krzywinski M.I., Skalka U., Smailus D.E., Schnerch A., Schein J.E.,
 RA Jones S.J., Marra M.A.,
 RT "Generation and initial analysis of more than 15,000 full-length human
 RT and mouse cDNA sequences.";
 RL Proc. Natl. Acad. Sci. U.S.A. 99:16899-16903(2002).
 RN [2]

RN SEQUENCE FROM N.A.
 RC STRAIN=AB; TISSUE=Whole body;
 RA Director MGC Project;
 RL Submitted (SEP-2003) to the EMBL/GenBank/DBJ databases.
 DR EMBL; BC057491; AAH57491.1; -;
 DR ZFIN; ZDB-GENE-030616-55; slc35a5.
 DR GO; GO:0000139; C:Golgi membrane; IEA.
 DR GO; GO:0016021; C:integral to membrane; IEA.
 DR GO; GO:0005338; F:nucleotide-sugar transporter activity; IEA.
 DR GO; GO:0005351; F:sugar porter activity; IEA.
 DR GO; GO:0015780; P:nucleotide-sugar transport; IEA.
 DR InterPro; IPR007271; Nuc_sug_transp.
 DR InterPro; IPR004689; UDPgal_transp.
 DR Pfam; PF04142; Nuc_sug_transp; 1.
 DR TIGRFAMs; TIGR00803; nst; 1.
 SQ SEQUENCE 436 AA; 48988 MW; D780C5D739AB1DFB CRC64;

Query Match 82.2%; Score 48.5; DB 2; Length 436;
 Best Local Similarity 91.7%; Pred. No. 53;
 Matches 11; Conservative 0; Mismatches 0; Indels 1; Gaps 1;

QY 1 SSSSSSSSCC-CC 11
 |||||
 DB 8 SSSSSSSSCGCC 19

RESULT 14
 Q9LKK0 PRELIMINARY; PRT; 240 AA.
 AC Q9LKK0;
 DT 01-OCT-2000 (TrEMBLrel. 15, Created)
 DT 01-OCT-2000 (TrEMBLrel. 15, Last sequence update)
 DT 01-MAR-2004 (TrEMBLrel. 26, Last annotation update)
 DE Apetala2 domain-containing protein.
 OS Atriplex hortensis (Mountain spinach).
 OC Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
 OC Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots;
 OC Caryophyllales; Amaranthaceae; Atriplex.
 OC NCBI_TaxID=34272;
 RN [1]
 RN SEQUENCE FROM N.A.
 RA Shen Y., Chen S.;
 RL Submitted (MAY-2000) to the EMBL/GenBank/DBJ databases.
 DR EMBL; AF274033; AAF76898.1; -;
 DR HSSP; O80337; 2GCC.
 DR GO; GO:0005634; C:nucleus; IEA.
 DR GO; GO:0003700; F:transcription factor activity; IEA.
 DR GO; GO:0006355; P:regulation of transcription, DNA-dependent; IEA.
 DR InterPro; IPR001471; TF_ERF.
 DR PRINTS; PR00367; ETHRSPFELMNT.
 DR PRODOM; PD001423; TF_ERF; 1.
 DR SMART; SM00380; AP2; 1.
 SQ SEQUENCE 240 AA; 26090 MW; A49764B2FFB49674 CRC64;

Query Match 79.7%; Score 47; DB 2; Length 240;
 Best Local Similarity 90.0%; Pred. No. 47;
 Matches 9; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

QY 1 SSSSSSSSCC 10
 :|||||
 DB 154 TSSSSSSSCC 163

RESULT 15
 Q6AV38 PRELIMINARY; PRT; 367 AA.
 AC Q6AV38;
 DT 25-OCT-2004 (TrEMBLrel. 28, Created)
 DT 25-OCT-2004 (TrEMBLrel. 28, Last sequence update)
 DT 25-OCT-2004 (TrEMBLrel. 28, Last annotation update)
 DE Putative Cyclin.
 GN Name=OSJNBa0063J18.4;
 OS Oryza sativa (japonica cultivar-group).
 OC Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
 OC Spermatophyta; Magnoliophyta; Liliopsida; Poales; Poaceae;
 OC Ehrhartoideae; Oryzaceae; Oryza.
 OC NCBI_TaxID=39947;
 RN [1]
 RN SEQUENCE FROM N.A.
 RA Buell C.R., Yuan Q., Qiyang S., Liu J., Ganeberger K., Jones K.M.,
 RA Overton II L.L., Tsitrin T., Kim M.M., Bera J.J., Jin S.S.,
 RA Padrosh D.W., Tallon L.J., Koo H., Ziemann V., Heiso J., Blunt S.,
 RA Vanaken S.S., Riedmuller S.B., Utterback T.T., Feldblyum T.V.,
 RA Yang Q.Q., Haas B.J., Suh B.B., Peterson J.J., Quackenbush J.,
 RA White O., Salzberg S.L., Fraser C.M.;
 RT "Oryza sativa chromosome 3 BAC OSJNBa0063J18 genomic sequence.";
 RL Submitted (JAN-2002) to the EMBL/GenBank/DBJ databases.
 RN [2]

RN SEQUENCE FROM N.A.
 RA Buell R.;
 RL Submitted (JUL-2004) to the EMBL/GenBank/DBJ databases.
 CC -!- SIMILARITY: Belongs to the cyclin family.
 DR EMBL; AC107206; AAT77041.1; -;
 DR GO; GO:0005634; C:nucleus; IEA.
 DR GO; GO:0000074; P:regulation of cell cycle; IEA.
 DR InterPro; IPR006670; Cyclin.
 DR InterPro; IPR004367; Cyclin_Cterm.
 DR InterPro; IPR011028; Cyclin_like.
 DR InterPro; IPR006671; Cyclin_N.
 DR Pfam; PF02984; Cyclin_C; 1.
 DR Pfam; PF0134; Cyclin_N; 1.
 DR SMART; SM00385; CYCLIN; 1.
 KW Cyclin.
 SQ SEQUENCE 367 AA; 38727 MW; 023BF8C0CDAD1D79 CRC64;

Query Match 79.7%; Score 47; DB 2; Length 367;
 Best Local Similarity 90.0%; Pred. No. 71;
 Matches 9; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

QY 1 SSSSSSSSCC 10
 |||||
 DB 78 SSSSSSSSCC 87

Search completed: November 15, 2005, 12:28:18
 Job time : 82.1842 secs

This Page Blank (uspto)

GenCore version 5.1.6
Copyright (c) 1993 - 2005 Compugen Ltd.

OM protein - protein search, using sw model

Run on: November 15, 2005, 12:07:58 ; Search time 87.1316 Seconds
(without alignments)
48.827 Million cell updates/sec

Title: US-10-820-998-4
Perfect score: 59
Sequence: 1 SSSSSSSSSCCC 11

Scoring table: BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 2105692 seqs, 386760381 residues

Total number of hits satisfying chosen parameters: 2105692

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : A_Geneseq_16Dec04:.*
1: Geneseqp1980s:.*
2: Geneseqp1980s:.*
3: Geneseqp2000s:.*
4: Geneseqp2001s:.*
5: Geneseqp2002s:.*
6: Geneseqp2003as:.*
7: Geneseqp2003bs:.*
8: Geneseqp2004s:.*

Pred. No. is the number of results predicted by chance to have a
score greater than or equal to the score of the result being printed,
and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	50	84.7	101	3 AAG20621	Aag20621 Arabidops
2	50	84.7	105	3 AAG56340	Aag56340 Arabidops
3	50	84.7	106	3 AAG58593	Aag58593 Arabidops
4	50	84.7	107	3 AAG60179	Aag60179 Arabidops
5	50	84.7	159	4 AAG20837	Abg20837 Novel hum
6	50	84.7	974	2 AAG60039	Aaw60039 Cotton ce
7	50	84.7	974	2 AAW73308	Aaw73308 Cellulose
8	48	81.4	273	7 ABO81687	Abo81687 Pseudomon
9	47	79.7	240	8 ADK68328	Adk68328 Plant AP2
10	47	79.7	240	8 ADI42828	Adi42828 Plant tra
11	47	79.7	238	8 ADP30061	Adp30061 Human sec
12	47	79.7	958	5 ABB92981	Abb92981 Herbicida
13	46	78.0	145	7 ADM06120	Adm06120 Human pro
14	46	78.0	216	7 ABO80411	Abo80411 Pseudomon
15	46	78.0	216	1 AAP93703	Aap93703 Sequence
16	46	78.0	216	2 AAW31582	Aaw31582 Eimeria t
17	46	78.0	216	7 ADF77354	Adf77354 20kd prot
18	46	78.0	237	1 AAP90502	Aap90502 Gp.B Eime
19	46	78.0	237	1 AAP90117	Aap90117 Group B i
20	46	78.0	237	1 AAW72600	Aaw72600 Group B E
21	46	78.0	248	2 AAR05221	Aar05221 Antigen G
22	45	76.3	66	4 AAU55134	Aau55134 Propionib
23	45	76.3	66	6 ABM51653	Abm51653 Propionib
24	45	76.3	83	8 ABO57055	Abo57055 Human gen
25	45	76.3	110	3 AAB42229	Aab42229 Human ORF

26	45	76.3	113	6 ABM65929	Abm65929 Propionib
27	45	76.3	212	4 ABG04843	Abg04843 Novel hum
28	45	76.3	402	4 ABG06268	Abg06268 Novel hum
29	45	76.3	462	4 AAU30404	Aau30404 Novel hum
30	45	76.3	895	6 ABU12304	Abu12304 Human PAT
31	45	76.3	933	6 ABU12307	Abu12307 Human PAT
32	45	76.3	969	6 ABU12301	Abu12301 Human PAT
33	45	76.3	1004	6 ABU12313	Abu12313 Human PAT
34	45	76.3	1044	6 ABU12310	Abu12310 Human PAT
35	44	74.6	74	2 AAY59967	Aay59967 Human end
36	44	74.6	99	5 ABP43926	Abp43926 NICE-1 pr
37	44	74.6	99	8 ADN04970	Adn04970 Antipsori
38	44	74.6	143	7 ABO68957	Abo68957 Pseudomon
39	44	74.6	322	4 ABG15587	Abg15587 Novel hum
40	44	74.6	327	7 ABO81872	Abo81872 Pseudomon
41	44	74.6	606	4 ABG15588	Abg15588 Novel hum
42	44	74.6	806	4 ABG07472	Abg07472 Novel hum
43	44	74.6	806	4 ABG15589	Abg15589 Novel hum
44	44	74.6	847	5 ABB76164	Abb76164 Human pot
45	43	72.9	58	4 AAU51336	Aau51336 Propionib

ALIGNMENTS

RESULT 1
AAG20621
ID AAG20621 standard; protein; 101 AA.
XX
AC AAG20621;

XX
DT 17-OCT-2000 (first entry)
XX
DE Arabidopsis thaliana protein fragment SEQ ID NO: 22884.
XX
KW Protein identification; signal transduction pathway; metabolic pathway;
KW hybridisation assay; genetic mapping; gene expression control; promoter;
KW termination sequence.

XX
OS Arabidopsis thaliana.
XX
PN EP1033405-A2.
XX
PD 06-SEP-2000.

XX
PF 25-FEB-2000; 2000EP-00301439.
XX
PR 25-FEB-1999; 99US-0121825P.
PR 05-MAR-1999; 99US-0123180P.
PR 09-MAR-1999; 99US-0123548P.
PR 23-MAR-1999; 99US-0125788P.
PR 25-MAR-1999; 99US-0126264P.
PR 29-MAR-1999; 99US-0126785P.
PR 01-APR-1999; 99US-0127462P.
PR 06-APR-1999; 99US-0128234P.
PR 08-APR-1999; 99US-0128714P.
PR 16-APR-1999; 99US-0129845P.
PR 19-APR-1999; 99US-0130077P.
PR 21-APR-1999; 99US-0130449P.
PR 23-APR-1999; 99US-0130510P.
PR 23-APR-1999; 99US-0130891P.
PR 28-APR-1999; 99US-0131449P.
PR 30-APR-1999; 99US-0132048P.
PR 04-MAY-1999; 99US-0132407P.
PR 05-MAY-1999; 99US-0132484P.
PR 06-MAY-1999; 99US-0132485P.
PR 06-MAY-1999; 99US-0132486P.
PR 07-MAY-1999; 99US-0132487P.
PR 11-MAY-1999; 99US-0132863P.
PR 14-MAY-1999; 99US-0134256P.
PR 14-MAY-1999; 99US-0134218P.
PR 14-MAY-1999; 99US-0134219P.
PR 14-MAY-1999; 99US-0134221P.

PR 14-MAY-1999; 99US-0134370P.
PR 18-MAY-1999; 99US-0134768P.
PR 19-MAY-1999; 99US-0134941P.
PR 20-MAY-1999; 99US-0135124P.
PR 21-MAY-1999; 99US-0135353P.
PR 24-MAY-1999; 99US-0135629P.
PR 25-MAY-1999; 99US-0136021P.
PR 27-MAY-1999; 99US-0136392P.
PR 28-MAY-1999; 99US-0136782P.
PR 01-JUN-1999; 99US-0137222P.
PR 03-JUN-1999; 99US-0137528P.
PR 04-JUN-1999; 99US-0137502P.
PR 07-JUN-1999; 99US-0137724P.
PR 08-JUN-1999; 99US-0138094P.
PR 10-JUN-1999; 99US-0138540P.
PR 10-JUN-1999; 99US-0138847P.
PR 14-JUN-1999; 99US-0139112P.
PR 16-JUN-1999; 99US-0139452P.
PR 16-JUN-1999; 99US-0139453P.
PR 17-JUN-1999; 99US-0139492P.
PR 18-JUN-1999; 99US-0139454P.
PR 18-JUN-1999; 99US-0139455P.
PR 18-JUN-1999; 99US-0139456P.
PR 18-JUN-1999; 99US-0139457P.
PR 18-JUN-1999; 99US-0139458P.
PR 18-JUN-1999; 99US-0139459P.
PR 18-JUN-1999; 99US-0139460P.
PR 18-JUN-1999; 99US-0139461P.
PR 18-JUN-1999; 99US-0139462P.
PR 18-JUN-1999; 99US-0139463P.
PR 18-JUN-1999; 99US-0139750P.
PR 18-JUN-1999; 99US-0139763P.
PR 21-JUN-1999; 99US-0139817P.
PR 22-JUN-1999; 99US-0139899P.
PR 23-JUN-1999; 99US-0140353P.
PR 23-JUN-1999; 99US-0140354P.
PR 24-JUN-1999; 99US-0140695P.
PR 28-JUN-1999; 99US-0140823P.
PR 29-JUN-1999; 99US-0140991P.
PR 30-JUN-1999; 99US-0141287P.
PR 01-JUL-1999; 99US-0141842P.
PR 01-JUL-1999; 99US-0142154P.
PR 02-JUL-1999; 99US-0142055P.
PR 06-JUL-1999; 99US-0142390P.
PR 08-JUL-1999; 99US-0142803P.
PR 09-JUL-1999; 99US-0142920P.
PR 12-JUL-1999; 99US-0142977P.
PR 13-JUL-1999; 99US-0143342P.
PR 14-JUL-1999; 99US-0143624P.
PR 15-JUL-1999; 99US-0144005P.
PR 16-JUL-1999; 99US-0144085P.
PR 16-JUL-1999; 99US-0144086P.
PR 19-JUL-1999; 99US-0144325P.
PR 19-JUL-1999; 99US-0144331P.
PR 19-JUL-1999; 99US-0144332P.
PR 19-JUL-1999; 99US-0144333P.
PR 19-JUL-1999; 99US-0144334P.
PR 19-JUL-1999; 99US-0144335P.
PR 20-JUL-1999; 99US-0144352P.
PR 20-JUL-1999; 99US-0144632P.
PR 20-JUL-1999; 99US-0144884P.
PR 21-JUL-1999; 99US-0144814P.
PR 21-JUL-1999; 99US-0145086P.
PR 21-JUL-1999; 99US-0145088P.
PR 22-JUL-1999; 99US-0145085P.
PR 22-JUL-1999; 99US-0145087P.
PR 22-JUL-1999; 99US-0145089P.
PR 22-JUL-1999; 99US-0145192P.
PR 23-JUL-1999; 99US-0145145P.
PR 23-JUL-1999; 99US-0145218P.
PR 23-JUL-1999; 99US-0145224P.
PR 26-JUL-1999; 99US-0145276P.
PR 27-JUL-1999; 99US-0145913P.
PR 27-JUL-1999; 99US-0134370P.
PR 27-JUL-1999; 99US-0134768P.
PR 28-JUL-1999; 99US-0134941P.
PR 02-AUG-1999; 99US-0135124P.
PR 02-AUG-1999; 99US-0135353P.
PR 02-AUG-1999; 99US-0135629P.
PR 03-AUG-1999; 99US-0136021P.
PR 04-AUG-1999; 99US-0136392P.
PR 04-AUG-1999; 99US-0136782P.
PR 05-AUG-1999; 99US-0137222P.
PR 05-AUG-1999; 99US-0137528P.
PR 06-AUG-1999; 99US-0137502P.
PR 06-AUG-1999; 99US-0137724P.
PR 09-AUG-1999; 99US-0138094P.
PR 09-AUG-1999; 99US-0138540P.
PR 10-AUG-1999; 99US-0138847P.
PR 11-AUG-1999; 99US-0139112P.
PR 12-AUG-1999; 99US-0139452P.
PR 13-AUG-1999; 99US-0139453P.
PR 13-AUG-1999; 99US-0139492P.
PR 16-AUG-1999; 99US-0139454P.
PR 16-AUG-1999; 99US-0139455P.
PR 16-AUG-1999; 99US-0139456P.
PR 20-AUG-1999; 99US-0139457P.
PR 20-AUG-1999; 99US-0139458P.
PR 20-AUG-1999; 99US-0139459P.
PR 23-AUG-1999; 99US-0139460P.
PR 23-AUG-1999; 99US-0139461P.
PR 25-AUG-1999; 99US-0139462P.
PR 26-AUG-1999; 99US-0139463P.
PR 27-AUG-1999; 99US-0139750P.
PR 27-AUG-1999; 99US-0139763P.
PR 30-AUG-1999; 99US-0139817P.
PR 31-AUG-1999; 99US-0139899P.
PR 01-SEP-1999; 99US-0140353P.
PR 01-SEP-1999; 99US-0140354P.
PR 07-SEP-1999; 99US-0140695P.
PR 10-SEP-1999; 99US-0140823P.
PR 13-SEP-1999; 99US-0140991P.
PR 15-SEP-1999; 99US-0141287P.
PR 16-SEP-1999; 99US-0141842P.
PR 20-SEP-1999; 99US-0142154P.
PR 22-SEP-1999; 99US-0142055P.
PR 23-SEP-1999; 99US-0142390P.
PR 24-SEP-1999; 99US-0142803P.
PR 28-SEP-1999; 99US-0142920P.
PR 29-SEP-1999; 99US-0142977P.
PR 04-OCT-1999; 99US-0143342P.
PR 05-OCT-1999; 99US-0143624P.
PR 06-OCT-1999; 99US-0144005P.
PR 07-OCT-1999; 99US-0144085P.
PR 08-OCT-1999; 99US-0144086P.
PR 12-OCT-1999; 99US-0144325P.
PR 13-OCT-1999; 99US-0144331P.
PR 13-OCT-1999; 99US-0144332P.
PR 13-OCT-1999; 99US-0144333P.
PR 14-OCT-1999; 99US-0144334P.
PR 14-OCT-1999; 99US-0144335P.
PR 14-OCT-1999; 99US-0144352P.
PR 14-OCT-1999; 99US-0144632P.
PR 18-OCT-1999; 99US-0144884P.
PR 18-OCT-1999; 99US-0144814P.
PR 21-OCT-1999; 99US-0145086P.
PR 21-OCT-1999; 99US-0145088P.
PR 21-OCT-1999; 99US-0145085P.
PR 21-OCT-1999; 99US-0145087P.
PR 22-OCT-1999; 99US-0145089P.
PR 22-OCT-1999; 99US-0145192P.
PR 23-OCT-1999; 99US-0145145P.
PR 23-OCT-1999; 99US-0145218P.
PR 23-OCT-1999; 99US-0145224P.
PR 25-OCT-1999; 99US-0145276P.
PR 25-OCT-1999; 99US-0145913P.
PR 27-JUL-1999; 99US-0145918P.
PR 27-JUL-1999; 99US-0145919P.
PR 28-JUL-1999; 99US-0145951P.
PR 02-AUG-1999; 99US-0146386P.
PR 02-AUG-1999; 99US-0146388P.
PR 02-AUG-1999; 99US-0146389P.
PR 03-AUG-1999; 99US-0147038P.
PR 04-AUG-1999; 99US-0147204P.
PR 04-AUG-1999; 99US-0147302P.
PR 05-AUG-1999; 99US-0147192P.
PR 05-AUG-1999; 99US-0147260P.
PR 06-AUG-1999; 99US-0147303P.
PR 06-AUG-1999; 99US-0147416P.
PR 09-AUG-1999; 99US-0147493P.
PR 09-AUG-1999; 99US-0147935P.
PR 10-AUG-1999; 99US-0148171P.
PR 11-AUG-1999; 99US-0148319P.
PR 12-AUG-1999; 99US-0148341P.
PR 13-AUG-1999; 99US-0148565P.
PR 13-AUG-1999; 99US-0148684P.
PR 16-AUG-1999; 99US-0149368P.
PR 16-AUG-1999; 99US-0149175P.
PR 18-AUG-1999; 99US-0149426P.
PR 20-AUG-1999; 99US-0149722P.
PR 20-AUG-1999; 99US-0149723P.
PR 20-AUG-1999; 99US-0149929P.
PR 23-AUG-1999; 99US-0149902P.
PR 23-AUG-1999; 99US-0149930P.
PR 25-AUG-1999; 99US-0150566P.
PR 26-AUG-1999; 99US-0150884P.
PR 27-AUG-1999; 99US-0151065P.
PR 27-AUG-1999; 99US-0151066P.
PR 30-AUG-1999; 99US-0151080P.
PR 31-AUG-1999; 99US-0151303P.
PR 01-SEP-1999; 99US-0151438P.
PR 01-SEP-1999; 99US-0151930P.
PR 10-SEP-1999; 99US-0152363P.
PR 10-SEP-1999; 99US-0153070P.
PR 13-SEP-1999; 99US-0153758P.
PR 15-SEP-1999; 99US-0154018P.
PR 16-SEP-1999; 99US-0154039P.
PR 20-SEP-1999; 99US-0154779P.
PR 22-SEP-1999; 99US-0155139P.
PR 23-SEP-1999; 99US-0155486P.
PR 24-SEP-1999; 99US-015659P.
PR 28-SEP-1999; 99US-0156458P.
PR 29-SEP-1999; 99US-0156596P.
PR 04-OCT-1999; 99US-0157117P.
PR 05-OCT-1999; 99US-0157753P.
PR 06-OCT-1999; 99US-0157865P.
PR 07-OCT-1999; 99US-0158029P.
PR 08-OCT-1999; 99US-0158232P.
PR 12-OCT-1999; 99US-0158369P.
PR 13-OCT-1999; 99US-0159293P.
PR 13-OCT-1999; 99US-0159294P.
PR 13-OCT-1999; 99US-0159295P.
PR 14-OCT-1999; 99US-0159329P.
PR 14-OCT-1999; 99US-0159330P.
PR 14-OCT-1999; 99US-0159331P.
PR 14-OCT-1999; 99US-0159637P.
PR 18-OCT-1999; 99US-0159638P.
PR 18-OCT-1999; 99US-0159584P.
PR 21-OCT-1999; 99US-0160741P.
PR 21-OCT-1999; 99US-0160767P.
PR 21-OCT-1999; 99US-0160768P.
PR 21-OCT-1999; 99US-0160770P.
PR 21-OCT-1999; 99US-0160814P.
PR 21-OCT-1999; 99US-0160815P.
PR 22-OCT-1999; 99US-0160980P.
PR 22-OCT-1999; 99US-0160981P.
PR 22-OCT-1999; 99US-0160989P.
PR 23-OCT-1999; 99US-0161404P.
PR 25-OCT-1999; 99US-0161405P.


```
PR 25-OCT-1999; 99US-0161406P.
PR 26-OCT-1999; 99US-0161355P.
PR 26-OCT-1999; 99US-0161360P.
PR 26-OCT-1999; 99US-0161361P.
PR 28-OCT-1999; 99US-0161920P.
PR 28-OCT-1999; 99US-0161922P.
PR 28-OCT-1999; 99US-0161933P.
PR 28-OCT-1999; 99US-0162142P.
PR 29-OCT-1999; 99US-0162142P.

Query Match      84.7%; Score 50; DB 3; Length 101;
Best Local Similarity 72.7%; Pred. No. 47;
Matches 8; Conservative 3; Mismatches 0; Indels 0; Gaps 0;

Qy 1 SSSSSSSSSCCC 11
Db 91 SASTSSSACCC 101

RESULT 2
AAG56340
ID AAG56340 standard; protein; 105 AA.
XX
AC AAG56340;
XX
DT 18-OCT-2000 (first entry)
XX
DE Arabidopsis thaliana protein fragment SEQ ID NO: 72406.
XX
KW Protein identification; signal transduction pathway; metabolic pathway;
KW hybridisation assay; genetic mapping; gene expression control; promoter;
KW termination sequence.
XX
OS Arabidopsis thaliana.
XX
PN EP1033405-A2.
XX
PD 06-SEP-2000.
XX
XX
XX 25-FEB-2000; 2000EP-00301439.
XX
PR 25-FEB-1999; 99US-0121825P.
PR 05-MAR-1999; 99US-0123180P.
PR 09-MAR-1999; 99US-0123548P.
PR 23-MAR-1999; 99US-0125788P.
PR 25-MAR-1999; 99US-0126264P.
PR 29-MAR-1999; 99US-0126785P.
PR 01-APR-1999; 99US-0127462P.
PR 08-APR-1999; 99US-0128234P.
PR 08-APR-1999; 99US-0128714P.
PR 16-APR-1999; 99US-0129845P.
PR 19-APR-1999; 99US-0130077P.
PR 21-APR-1999; 99US-0130449P.
PR 23-APR-1999; 99US-0130510P.
PR 23-APR-1999; 99US-0130891P.
PR 28-APR-1999; 99US-0131449P.
PR 30-APR-1999; 99US-0132048P.
PR 30-APR-1999; 99US-0132407P.
PR 04-MAY-1999; 99US-0132484P.
PR 05-MAY-1999; 99US-0132485P.
PR 06-MAY-1999; 99US-0132486P.
PR 06-MAY-1999; 99US-0132487P.
PR 07-MAY-1999; 99US-0132863P.
PR 11-MAY-1999; 99US-0134256P.
PR 14-MAY-1999; 99US-0134218P.
PR 14-MAY-1999; 99US-0134219P.
PR 14-MAY-1999; 99US-0134221P.
PR 14-MAY-1999; 99US-0134370P.
PR 18-MAY-1999; 99US-0134768P.
PR 19-MAY-1999; 99US-0134941P.
PR 20-MAY-1999; 99US-0135124P.
PR 21-MAY-1999; 99US-0135553P.
PR 24-MAY-1999; 99US-0135629P.
PR 25-MAY-1999; 99US-0136021P.
PR
PR 25-OCT-1999; 99US-0136392P.
PR 28-MAY-1999; 99US-0136782P.
PR 01-JUN-1999; 99US-0137222P.
PR 03-JUN-1999; 99US-0137528P.
PR 04-JUN-1999; 99US-0137502P.
PR 07-JUN-1999; 99US-0137724P.
PR 08-JUN-1999; 99US-0138094P.
PR 10-JUN-1999; 99US-0138540P.
PR 10-JUN-1999; 99US-0138847P.
PR 14-JUN-1999; 99US-0139119P.
PR 16-JUN-1999; 99US-0139452P.
PR 16-JUN-1999; 99US-0139453P.
PR 17-JUN-1999; 99US-0139492P.
PR 18-JUN-1999; 99US-0139454P.
PR 18-JUN-1999; 99US-0139455P.
PR 18-JUN-1999; 99US-0139456P.
PR 18-JUN-1999; 99US-0139457P.
PR 18-JUN-1999; 99US-0139458P.
PR 18-JUN-1999; 99US-0139460P.
PR 18-JUN-1999; 99US-0139461P.
PR 18-JUN-1999; 99US-0139462P.
PR 18-JUN-1999; 99US-0139463P.
PR 18-JUN-1999; 99US-0139750P.
PR 18-JUN-1999; 99US-0139763P.
PR 21-JUN-1999; 99US-0139817P.
PR 22-JUN-1999; 99US-0139899P.
PR 23-JUN-1999; 99US-0140353P.
PR 23-JUN-1999; 99US-0140354P.
PR 24-JUN-1999; 99US-0140695P.
PR 28-JUN-1999; 99US-0140823P.
PR 29-JUN-1999; 99US-0140991P.
PR 30-JUN-1999; 99US-0141287P.
PR 01-JUL-1999; 99US-0141842P.
PR 02-JUL-1999; 99US-0142154P.
PR 02-JUL-1999; 99US-0142055P.
PR 06-JUL-1999; 99US-0142390P.
PR 08-JUL-1999; 99US-0142803P.
PR 09-JUL-1999; 99US-0142920P.
PR 12-JUL-1999; 99US-0142977P.
PR 13-JUL-1999; 99US-0143542P.
PR 14-JUL-1999; 99US-0143624P.
PR 15-JUL-1999; 99US-0144005P.
PR 16-JUL-1999; 99US-0144085P.
PR 16-JUL-1999; 99US-0144086P.
PR 19-JUL-1999; 99US-0144325P.
PR 19-JUL-1999; 99US-0144331P.
PR 19-JUL-1999; 99US-0144332P.
PR 19-JUL-1999; 99US-0144333P.
PR 19-JUL-1999; 99US-0144334P.
PR 19-JUL-1999; 99US-0144335P.
PR 20-JUL-1999; 99US-0144352P.
PR 20-JUL-1999; 99US-0144632P.
PR 20-JUL-1999; 99US-0144884P.
PR 21-JUL-1999; 99US-0144814P.
PR 21-JUL-1999; 99US-0145086P.
PR 21-JUL-1999; 99US-0145088P.
PR 22-JUL-1999; 99US-0145085P.
PR 22-JUL-1999; 99US-0145087P.
PR 22-JUL-1999; 99US-0145089P.
PR 22-JUL-1999; 99US-0145192P.
PR 23-JUL-1999; 99US-0145145P.
PR 23-JUL-1999; 99US-0145218P.
PR 23-JUL-1999; 99US-0145224P.
PR 26-JUL-1999; 99US-0145276P.
PR 27-JUL-1999; 99US-0145513P.
PR 27-JUL-1999; 99US-0145918P.
PR 27-JUL-1999; 99US-0145919P.
PR 28-JUL-1999; 99US-0145951P.
PR 02-AUG-1999; 99US-0146386P.
PR 02-AUG-1999; 99US-0146388P.
PR 02-AUG-1999; 99US-0146389P.
PR 03-AUG-1999; 99US-0147038P.
```

```
PR 04-AUG-1999; 99US-0147204P.
PR 05-AUG-1999; 99US-0147302P.
PR 05-AUG-1999; 99US-0147192P.
PR 05-AUG-1999; 99US-0147260P.
PR 06-AUG-1999; 99US-0147303P.
PR 06-AUG-1999; 99US-0147416P.
PR 09-AUG-1999; 99US-0147493P.
PR 09-AUG-1999; 99US-0147935P.
PR 10-AUG-1999; 99US-0148171P.
PR 11-AUG-1999; 99US-0148319P.
PR 12-AUG-1999; 99US-0148341P.
PR 13-AUG-1999; 99US-0148565P.
PR 13-AUG-1999; 99US-0148684P.
PR 16-AUG-1999; 99US-0149368P.
PR 17-AUG-1999; 99US-0149175P.
PR 18-AUG-1999; 99US-0149426P.
PR 20-AUG-1999; 99US-0149722P.
PR 20-AUG-1999; 99US-0149723P.
PR 20-AUG-1999; 99US-0149929P.
PR 23-AUG-1999; 99US-0149902P.
PR 23-AUG-1999; 99US-0149930P.
PR 25-AUG-1999; 99US-0150566P.
PR 28-AUG-1999; 99US-0150884P.
PR 27-AUG-1999; 99US-0151065P.
PR 27-AUG-1999; 99US-0151066P.
PR 27-AUG-1999; 99US-0151080P.
PR 30-AUG-1999; 99US-0151303P.
PR 31-AUG-1999; 99US-0151438P.
PR 01-SEP-1999; 99US-0151930P.
PR 07-SEP-1999; 99US-0152363P.
PR 10-SEP-1999; 99US-0153070P.
PR 13-SEP-1999; 99US-0153758P.
PR 15-SEP-1999; 99US-0154018P.
PR 16-SEP-1999; 99US-0154039P.
PR 20-SEP-1999; 99US-0154779P.
PR 22-SEP-1999; 99US-0155139P.
PR 23-SEP-1999; 99US-0155486P.
PR 24-SEP-1999; 99US-0155659P.
PR 28-SEP-1999; 99US-0156458P.
PR 29-SEP-1999; 99US-0156596P.
PR 04-OCT-1999; 99US-0157117P.
PR 05-OCT-1999; 99US-0157753P.
PR 06-OCT-1999; 99US-0158029P.
PR 07-OCT-1999; 99US-0158232P.
PR 08-OCT-1999; 99US-0158369P.
PR 12-OCT-1999; 99US-0159293P.
PR 13-OCT-1999; 99US-0159294P.
PR 13-OCT-1999; 99US-0159295P.
PR 14-OCT-1999; 99US-0159339P.
PR 14-OCT-1999; 99US-0159331P.
PR 14-OCT-1999; 99US-0159637P.
PR 14-OCT-1999; 99US-0159638P.
PR 18-OCT-1999; 99US-0159584P.
PR 21-OCT-1999; 99US-0160741P.
PR 21-OCT-1999; 99US-0160767P.
PR 21-OCT-1999; 99US-0160768P.
PR 21-OCT-1999; 99US-0160770P.
PR 21-OCT-1999; 99US-0160814P.
PR 21-OCT-1999; 99US-0160815P.
PR 22-OCT-1999; 99US-0160980P.
PR 22-OCT-1999; 99US-0160981P.
PR 22-OCT-1999; 99US-0160989P.
PR 25-OCT-1999; 99US-0161404P.
PR 25-OCT-1999; 99US-0161405P.
PR 25-OCT-1999; 99US-0161406P.
PR 26-OCT-1999; 99US-0161359P.
PR 26-OCT-1999; 99US-0161360P.
PR 26-OCT-1999; 99US-0161361P.
PR 28-OCT-1999; 99US-0161920P.
PR 28-OCT-1999; 99US-0161922P.
PR 28-OCT-1999; 99US-0161933P.

PR 29-OCT-1999; 99US-0162142P.
Query Match 84.7%; Score 50; DB 3; Length 105;
Best Local Similarity 100.0%; Pred. No. 48;
Matches 10; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 SSSSSSSSCC 10
Db 4 SSSSSSSSCC 13

RESULT 3
AAG58593
ID AAG58593 standard; protein; 106 AA.
XX AC AAG58593;
XX DT 18-OCT-2000 (first entry)
XX DE Arabidopsis thaliana protein fragment SEQ ID NO: 75659.
XX KW Protein identification; signal transduction pathway; metabolic pathway;
XX KW hybridisation assay; genetic mapping; gene expression control; promoter;
XX OS Arabidopsis thaliana.
XX PN EP1033405-A2.
XX PD 06-SEP-2000.
XX PF 25-FEB-2000; 2000EP-00301439.
XX PR 25-FEB-1999; 99US-0121825P.
PR 05-MAR-1999; 99US-0123180P.
PR 09-MAR-1999; 99US-0123548P.
PR 23-MAR-1999; 99US-0125788P.
PR 25-MAR-1999; 99US-0126264P.
PR 29-MAR-1999; 99US-0126785P.
PR 01-APR-1999; 99US-0127462P.
PR 06-APR-1999; 99US-0128234P.
PR 08-APR-1999; 99US-0128714P.
PR 16-APR-1999; 99US-0129845P.
PR 19-APR-1999; 99US-0130077P.
PR 21-APR-1999; 99US-0130449P.
PR 23-APR-1999; 99US-0130510P.
PR 28-APR-1999; 99US-0131449P.
PR 30-APR-1999; 99US-0132048P.
PR 30-APR-1999; 99US-0132407P.
PR 04-MAY-1999; 99US-0132484P.
PR 05-MAY-1999; 99US-0132485P.
PR 06-MAY-1999; 99US-0132486P.
PR 06-MAY-1999; 99US-0132487P.
PR 11-MAY-1999; 99US-0132863P.
PR 17-MAY-1999; 99US-0134256P.
PR 14-MAY-1999; 99US-0134218P.
PR 14-MAY-1999; 99US-0134219P.
PR 14-MAY-1999; 99US-0134221P.
PR 14-MAY-1999; 99US-0134370P.
PR 18-MAY-1999; 99US-0134768P.
PR 19-MAY-1999; 99US-0134941P.
PR 20-MAY-1999; 99US-0135124P.
PR 21-MAY-1999; 99US-0135353P.
PR 24-MAY-1999; 99US-0135629P.
PR 25-MAY-1999; 99US-0136021P.
PR 27-MAY-1999; 99US-0136392P.
PR 28-MAY-1999; 99US-0136782P.
PR 01-JUN-1999; 99US-0137222P.
PR 03-JUN-1999; 99US-0137528P.
PR 04-JUN-1999; 99US-0137502P.
PR 07-JUN-1999; 99US-0137724P.
PR 08-JUN-1999; 99US-0138094P.
```

PR 10-JUN-1999; 99US-0138540P.
PR 10-JUN-1999; 99US-0138847P.
PR 14-JUN-1999; 99US-0139119P.
PR 16-JUN-1999; 99US-0139452P.
PR 16-JUN-1999; 99US-0139453P.
PR 17-JUN-1999; 99US-0139492P.
PR 18-JUN-1999; 99US-0139454P.
PR 18-JUN-1999; 99US-0139455P.
PR 18-JUN-1999; 99US-0139456P.
PR 18-JUN-1999; 99US-0139457P.
PR 18-JUN-1999; 99US-0139458P.
PR 18-JUN-1999; 99US-0139459P.
PR 18-JUN-1999; 99US-0139460P.
PR 18-JUN-1999; 99US-0139461P.
PR 18-JUN-1999; 99US-0139462P.
PR 18-JUN-1999; 99US-0139463P.
PR 18-JUN-1999; 99US-0139750P.
PR 18-JUN-1999; 99US-0139763P.
PR 21-JUN-1999; 99US-0139817P.
PR 22-JUN-1999; 99US-0139899P.
PR 23-JUN-1999; 99US-0140333P.
PR 23-JUN-1999; 99US-0140354P.
PR 24-JUN-1999; 99US-0140695P.
PR 28-JUN-1999; 99US-0140823P.
PR 28-JUN-1999; 99US-0140931P.
PR 30-JUN-1999; 99US-0141287P.
PR 01-JUL-1999; 99US-0141842P.
PR 01-JUL-1999; 99US-0142154P.
PR 02-JUL-1999; 99US-0142055P.
PR 06-JUL-1999; 99US-0142330P.
PR 08-JUL-1999; 99US-0142803P.
PR 08-JUL-1999; 99US-0142320P.
PR 12-JUL-1999; 99US-0142977P.
PR 13-JUL-1999; 99US-0143354P.
PR 14-JUL-1999; 99US-0143624P.
PR 15-JUL-1999; 99US-0144005P.
PR 16-JUL-1999; 99US-0144085P.
PR 16-JUL-1999; 99US-0144086P.
PR 19-JUL-1999; 99US-0144325P.
PR 19-JUL-1999; 99US-0144331P.
PR 19-JUL-1999; 99US-0144332P.
PR 19-JUL-1999; 99US-0144333P.
PR 19-JUL-1999; 99US-0144334P.
PR 19-JUL-1999; 99US-0144335P.
PR 20-JUL-1999; 99US-0144352P.
PR 20-JUL-1999; 99US-0144632P.
PR 20-JUL-1999; 99US-0144884P.
PR 21-JUL-1999; 99US-0144814P.
PR 21-JUL-1999; 99US-0145086P.
PR 21-JUL-1999; 99US-0145088P.
PR 22-JUL-1999; 99US-0145085P.
PR 22-JUL-1999; 99US-0145087P.
PR 22-JUL-1999; 99US-0145089P.
PR 22-JUL-1999; 99US-0145192P.
PR 23-JUL-1999; 99US-0145145P.
PR 23-JUL-1999; 99US-0145218P.
PR 23-JUL-1999; 99US-0145224P.
PR 26-JUL-1999; 99US-0145276P.
PR 27-JUL-1999; 99US-0145913P.
PR 27-JUL-1999; 99US-0145918P.
PR 27-JUL-1999; 99US-0145919P.
PR 28-JUL-1999; 99US-0145951P.
PR 28-JUL-1999; 99US-0146386P.
PR 02-AUG-1999; 99US-0146388P.
PR 02-AUG-1999; 99US-0146389P.
PR 03-AUG-1999; 99US-0147038P.
PR 04-AUG-1999; 99US-0147204P.
PR 04-AUG-1999; 99US-0147302P.
PR 05-AUG-1999; 99US-0147192P.
PR 05-AUG-1999; 99US-0147260P.
PR 06-AUG-1999; 99US-0147303P.
PR 06-AUG-1999; 99US-0147416P.
PR 09-AUG-1999; 99US-0147493P.

PR 09-AUG-1999; 99US-0147935P.
PR 10-AUG-1999; 99US-0148171P.
PR 11-AUG-1999; 99US-0148319P.
PR 12-AUG-1999; 99US-0148341P.
PR 13-AUG-1999; 99US-0148565P.
PR 13-AUG-1999; 99US-0148684P.
PR 16-AUG-1999; 99US-0149368P.
PR 17-AUG-1999; 99US-0149175P.
PR 18-AUG-1999; 99US-0149426P.
PR 20-AUG-1999; 99US-0149722P.
PR 20-AUG-1999; 99US-0149723P.
PR 20-AUG-1999; 99US-0149929P.
PR 23-AUG-1999; 99US-0149802P.
PR 23-AUG-1999; 99US-0149930P.
PR 25-AUG-1999; 99US-0150566P.
PR 26-AUG-1999; 99US-0150884P.
PR 27-AUG-1999; 99US-0151066P.
PR 27-AUG-1999; 99US-0151080P.
PR 30-AUG-1999; 99US-0151303P.
PR 31-AUG-1999; 99US-0151438P.
PR 01-SEP-1999; 99US-0151930P.
PR 07-SEP-1999; 99US-0152363P.
PR 10-SEP-1999; 99US-0153070P.
PR 13-SEP-1999; 99US-0153758P.
PR 15-SEP-1999; 99US-0154018P.
PR 16-SEP-1999; 99US-0154039P.
PR 20-SEP-1999; 99US-0154779P.
PR 22-SEP-1999; 99US-0155139P.
PR 23-SEP-1999; 99US-0155486P.
PR 24-SEP-1999; 99US-0155659P.
PR 28-SEP-1999; 99US-0156458P.
PR 29-SEP-1999; 99US-0156596P.
PR 04-OCT-1999; 99US-0157117P.
PR 05-OCT-1999; 99US-0157753P.
PR 06-OCT-1999; 99US-0157865P.
PR 07-OCT-1999; 99US-0158029P.
PR 08-OCT-1999; 99US-0158232P.
PR 12-OCT-1999; 99US-0158369P.
PR 13-OCT-1999; 99US-0159293P.
PR 13-OCT-1999; 99US-0159294P.
PR 14-OCT-1999; 99US-0159329P.
PR 14-OCT-1999; 99US-0159330P.
PR 14-OCT-1999; 99US-0159331P.
PR 14-OCT-1999; 99US-0159637P.
PR 18-OCT-1999; 99US-0159638P.
PR 18-OCT-1999; 99US-0159584P.
PR 21-OCT-1999; 99US-0160741P.
PR 21-OCT-1999; 99US-0160767P.
PR 21-OCT-1999; 99US-0160768P.
PR 21-OCT-1999; 99US-0160770P.
PR 21-OCT-1999; 99US-0160814P.
PR 21-OCT-1999; 99US-0160815P.
PR 22-OCT-1999; 99US-0160980P.
PR 22-OCT-1999; 99US-0160981P.
PR 22-OCT-1999; 99US-0160989P.
PR 25-OCT-1999; 99US-0161404P.
PR 25-OCT-1999; 99US-0161405P.
PR 25-OCT-1999; 99US-0161406P.
PR 26-OCT-1999; 99US-0161359P.
PR 26-OCT-1999; 99US-0161360P.
PR 28-OCT-1999; 99US-0161361P.
PR 28-OCT-1999; 99US-0161920P.
PR 28-OCT-1999; 99US-0161992P.
PR 28-OCT-1999; 99US-0161993P.
PR 29-OCT-1999; 99US-0162142P.

Query Match 84.7%; Score 50; DB 3; Length 106;
Best Local Similarity 100.0%; Pred. No. 49;
Matches 10; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 SSSSSSSSCC 10

```
Db          |||||||||
            5 SSSSSSSCC 14

RESULT 4
AAG60179
ID   AAG60179 standard; protein; 107 AA.
XX
AC   AAG60179;
XX
DT   18-OCT-2000 (first entry)
XX
DE   Arabidopsis thaliana protein fragment SEQ ID NO: 77918.
XX
KW   Protein identification; signal transduction pathway; metabolic pathway;
KW   hybridisation assay; genetic mapping; gene expression control; promoter;
KW   termination sequence.
XX
OS   Arabidopsis thaliana.
XX
PN   EP1033405-A2.
XX
PD   06-SEP-2000.
XX
XX
XX   25-FEB-2000; 2000EP-00301439.
XX
XX   25-FEB-1999; 99US-0121825P.
XX   05-MAR-1999; 99US-0123180P.
XX   09-MAR-1999; 99US-0123548P.
XX   23-MAR-1999; 99US-0125788P.
XX   25-MAR-1999; 99US-0126264P.
XX   29-MAR-1999; 99US-0126785P.
XX   01-APR-1999; 99US-0127452P.
XX   06-APR-1999; 99US-0128234P.
XX   08-APR-1999; 99US-0128714P.
XX   16-APR-1999; 99US-0129845P.
XX   19-APR-1999; 99US-0130077P.
XX   21-APR-1999; 99US-0130449P.
XX   23-APR-1999; 99US-0130510P.
XX   28-APR-1999; 99US-0131449P.
XX   30-APR-1999; 99US-0132048P.
XX   30-APR-1999; 99US-0132407P.
XX   04-MAY-1999; 99US-0132484P.
XX   05-MAY-1999; 99US-0132485P.
XX   06-MAY-1999; 99US-0132486P.
XX   07-MAY-1999; 99US-0132863P.
XX   11-MAY-1999; 99US-0134256P.
XX   14-MAY-1999; 99US-0134218P.
XX   14-MAY-1999; 99US-0134219P.
XX   14-MAY-1999; 99US-0134221P.
XX   14-MAY-1999; 99US-0134370P.
XX   18-MAY-1999; 99US-0134768P.
XX   19-MAY-1999; 99US-0134941P.
XX   20-MAY-1999; 99US-0135112P.
XX   21-MAY-1999; 99US-0135353P.
XX   24-MAY-1999; 99US-0135623P.
XX   25-MAY-1999; 99US-0136021P.
XX   27-MAY-1999; 99US-0136392P.
XX   28-MAY-1999; 99US-0136782P.
XX   01-JUN-1999; 99US-0137222P.
XX   03-JUN-1999; 99US-0137528P.
XX   04-JUN-1999; 99US-0137502P.
XX   07-JUN-1999; 99US-0137724P.
XX   08-JUN-1999; 99US-0138094P.
XX   10-JUN-1999; 99US-0138540P.
XX   10-JUN-1999; 99US-0138847P.
XX   14-JUN-1999; 99US-0139119P.
XX   16-JUN-1999; 99US-0139452P.
XX   16-JUN-1999; 99US-0139453P.
XX   17-JUN-1999; 99US-0139492P.
XX   18-JUN-1999; 99US-0139454P.
XX   18-JUN-1999; 99US-0139455P.
XX   18-JUN-1999; 99US-0139456P.
XX   18-JUN-1999; 99US-0139457P.
XX   18-JUN-1999; 99US-0139458P.
XX   18-JUN-1999; 99US-0139459P.
XX   18-JUN-1999; 99US-0139460P.
XX   18-JUN-1999; 99US-0139461P.
XX   18-JUN-1999; 99US-0139462P.
XX   18-JUN-1999; 99US-0139463P.
XX   18-JUN-1999; 99US-0139750P.
XX   18-JUN-1999; 99US-0139763P.
XX   21-JUN-1999; 99US-0139817P.
XX   22-JUN-1999; 99US-0139899P.
XX   23-JUN-1999; 99US-0140353P.
XX   23-JUN-1999; 99US-0140354P.
XX   24-JUN-1999; 99US-0140695P.
XX   28-JUN-1999; 99US-0140823P.
XX   29-JUN-1999; 99US-0140991P.
XX   30-JUN-1999; 99US-0141287P.
XX   01-JUL-1999; 99US-0141842P.
XX   02-JUL-1999; 99US-0142055P.
XX   06-JUL-1999; 99US-0142390P.
XX   08-JUL-1999; 99US-0142803P.
XX   09-JUL-1999; 99US-0142920P.
XX   12-JUL-1999; 99US-0142977P.
XX   13-JUL-1999; 99US-0143542P.
XX   14-JUL-1999; 99US-0143624P.
XX   15-JUL-1999; 99US-0144005P.
XX   16-JUL-1999; 99US-0144085P.
XX   16-JUL-1999; 99US-0144086P.
XX   19-JUL-1999; 99US-0144325P.
XX   19-JUL-1999; 99US-0144331P.
XX   19-JUL-1999; 99US-0144332P.
XX   19-JUL-1999; 99US-0144333P.
XX   19-JUL-1999; 99US-0144334P.
XX   19-JUL-1999; 99US-0144335P.
XX   20-JUL-1999; 99US-0144352P.
XX   20-JUL-1999; 99US-0144632P.
XX   20-JUL-1999; 99US-0144884P.
XX   21-JUL-1999; 99US-0144814P.
XX   21-JUL-1999; 99US-0145086P.
XX   21-JUL-1999; 99US-0145088P.
XX   22-JUL-1999; 99US-0145085P.
XX   22-JUL-1999; 99US-0145087P.
XX   22-JUL-1999; 99US-0145089P.
XX   22-JUL-1999; 99US-0145192P.
XX   23-JUL-1999; 99US-0145145P.
XX   23-JUL-1999; 99US-0145218P.
XX   23-JUL-1999; 99US-0145224P.
XX   26-JUL-1999; 99US-0145276P.
XX   27-JUL-1999; 99US-0145913P.
XX   27-JUL-1999; 99US-0145918P.
XX   27-JUL-1999; 99US-0145919P.
XX   28-JUL-1999; 99US-0145951P.
XX   02-AUG-1999; 99US-0146386P.
XX   02-AUG-1999; 99US-0146388P.
XX   02-AUG-1999; 99US-0146389P.
XX   03-AUG-1999; 99US-0147038P.
XX   04-AUG-1999; 99US-0147204P.
XX   04-AUG-1999; 99US-0147302P.
XX   05-AUG-1999; 99US-0147192P.
XX   05-AUG-1999; 99US-0147260P.
XX   06-AUG-1999; 99US-0147303P.
XX   06-AUG-1999; 99US-0147416P.
XX   09-AUG-1999; 99US-0147493P.
XX   09-AUG-1999; 99US-0147935P.
XX   10-AUG-1999; 99US-0148171P.
XX   11-AUG-1999; 99US-0148319P.
XX   12-AUG-1999; 99US-0148341P.
XX   13-AUG-1999; 99US-0148565P.
XX   13-AUG-1999; 99US-0148684P.
XX   16-AUG-1999; 99US-0149368P.
```

```
PR 17-AUG-1999; 99US-0149175P.
PR 18-AUG-1999; 99US-0149426P.
PR 20-AUG-1999; 99US-0149722P.
PR 20-AUG-1999; 99US-0149723P.
PR 20-AUG-1999; 99US-0149929P.
PR 23-AUG-1999; 99US-0149802P.
PR 23-AUG-1999; 99US-0149930P.
PR 25-AUG-1999; 99US-0150366P.
PR 26-AUG-1999; 99US-0150884P.
PR 27-AUG-1999; 99US-0151065P.
PR 27-AUG-1999; 99US-0151066P.
PR 27-AUG-1999; 99US-0151080P.
PR 30-AUG-1999; 99US-0151303P.
PR 31-AUG-1999; 99US-0151438P.
PR 01-SEP-1999; 99US-0151930P.
PR 07-SEP-1999; 99US-0152363P.
PR 10-SEP-1999; 99US-0153070P.
PR 13-SEP-1999; 99US-0153758P.
PR 15-SEP-1999; 99US-0154018P.
PR 16-SEP-1999; 99US-0154039P.
PR 20-SEP-1999; 99US-0154779P.
PR 22-SEP-1999; 99US-0155139P.
PR 23-SEP-1999; 99US-0155486P.
PR 24-SEP-1999; 99US-0155659P.
PR 28-SEP-1999; 99US-0156458P.
PR 29-SEP-1999; 99US-0156596P.
PR 04-OCT-1999; 99US-0157117P.
PR 05-OCT-1999; 99US-0157753P.
PR 06-OCT-1999; 99US-0157865P.
PR 07-OCT-1999; 99US-0158029P.
PR 08-OCT-1999; 99US-0158232P.
PR 12-OCT-1999; 99US-0158369P.
PR 13-OCT-1999; 99US-0159293P.
PR 13-OCT-1999; 99US-0159294P.
PR 13-OCT-1999; 99US-0159295P.
PR 14-OCT-1999; 99US-0159329P.
PR 14-OCT-1999; 99US-0159330P.
PR 14-OCT-1999; 99US-0159331P.
PR 14-OCT-1999; 99US-0159637P.
PR 14-OCT-1999; 99US-0159638P.
PR 18-OCT-1999; 99US-0159584P.
PR 21-OCT-1999; 99US-0160741P.
PR 21-OCT-1999; 99US-0160767P.
PR 21-OCT-1999; 99US-0160768P.
PR 21-OCT-1999; 99US-0160770P.
PR 21-OCT-1999; 99US-0160814P.
PR 21-OCT-1999; 99US-0160815P.
PR 22-OCT-1999; 99US-0160980P.
PR 22-OCT-1999; 99US-0160981P.
PR 22-OCT-1999; 99US-0160989P.
PR 25-OCT-1999; 99US-0161404P.
PR 25-OCT-1999; 99US-0161405P.
PR 25-OCT-1999; 99US-0161406P.
PR 26-OCT-1999; 99US-0161359P.
PR 26-OCT-1999; 99US-0161360P.
PR 26-OCT-1999; 99US-0161361P.
PR 28-OCT-1999; 99US-0161920P.
PR 28-OCT-1999; 99US-0161992P.
PR 28-OCT-1999; 99US-0161993P.
PR 29-OCT-1999; 99US-0162142P.

Query Match 84.7%; Score 50; DB 3; Length 107;
Best Local Similarity 100.0%; Pred. No. 49;
Matches 10; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 SSSSSSSSSCC 10
Db |||||

RESULT 5
ABG20837
ID ABG20837 standard; protein; 159 AA.

XX AC ABG20837;
XX 13-FEB-2002 (first entry)
XX DT Novel human diagnostic protein #20828.
XX DE Human; chromosome mapping; gene mapping; gene therapy; forensic;
XX KW food supplement; medical imaging; diagnostic; genetic disorder.
XX OS Homo sapiens.
XX PN WO200175067-A2.
XX PD 11-OCT-2001.
XX PF 30-MAR-2001; 2001WO-US008631.
XX PR 31-MAR-2000; 2000US-00540217.
XX PR 23-AUG-2000; 2000US-00649167.
XX PA (HYSE-) HYSEQ INC.
XX PI Drmanac RT, Liu C, Tang YT;
XX DR WPI; 2001-639362/73.
XX DR N-PSDB; AAS85024.
XX PT New isolated polynucleotide and encoded polypeptides, useful in
XX PT diagnostics, forensics, gene mapping, identification of mutations
XX PT responsible for genetic disorders or other traits and to assess
XX PT biodiversity.
XX PS Claim 20; SEQ ID NO 51196; 103pp; English.
XX CC The invention relates to isolated polynucleotide (I) and polypeptide (II)
XX CC sequences. (I) is useful as hybridisation probes, polymerase chain
XX CC reaction (PCR) primers, oligomers, and for chromosome and gene mapping,
XX CC and in recombinant production of (II). The polynucleotides are also used
XX CC in diagnostics as expressed sequence tags for identifying expressed
XX CC genes. (I) is useful in gene therapy techniques to restore normal
XX CC activity of (II) or to treat disease states involving (II). (II) is
XX CC useful for generating antibodies against it, detecting or quantitating a
XX CC polypeptide in tissue, as molecular weight markers and as a food
XX CC supplement. (II) and its binding partners are useful in medical imaging
XX CC of sites expressing (II). (I) and (II) are useful for treating disorders
XX CC involving aberrant protein expression or biological activities. The
XX CC polypeptide and polynucleotide sequences have applications in
XX CC diagnostics, forensics, gene mapping, identification of mutations
XX CC responsible for genetic disorders or other traits to assess biodiversity
XX CC and to produce other types of data and products dependent on DNA and
XX CC amino acid sequences. ABG00010-ABG30377 represent novel human diagnostic
XX CC amino acid sequences of the invention. Note: The sequence data for this
XX CC patent did not appear in the printed specification, but was obtained in
XX CC electronic format directly from WIPO at
XX CC ftp.wipo.int/pub/published_pct_sequences
XX SQ Sequence 159 AA;

Query Match 84.7%; Score 50; DB 4; Length 159;
Best Local Similarity 81.8%; Pred. No. 68;
Matches 9; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

QY 1 SSSSSSSSSCC 11
Db |||||

RESULT 6
AAW60039
ID AAW60039 standard; protein; 974 AA.
XX AC AAW60039;
```

XX 17-OCT-2003 (revised)
 DT 28-SEP-1998 (first entry)
 XX Cotton cellulose synthase CelA1.
 XX Cellulose synthase; CelA1; cotton; forest tree; plant fibre; wood;
 KW transgenic plant; ss.
 XX Gossypium hirsutum; Acala SJ-2.
 XX Key Location/Qualifiers
 FT Modified-site 81
 FT Region /note= "N-glycosylated"
 FT Region 175..295
 FT Region /label= H-1
 FT Region 280..295
 FT Region /label= U-1
 FT Misc-difference 293
 FT Region /note= "critical Asp residue"
 FT Region 296..420
 FT Region /label= P-CR
 FT Region 421..539
 FT Region /label= H-2
 FT Region 450..467
 FT Region /label= U-2
 FT Misc-difference 461
 FT Region /note= "critical Asp residue"
 FT Modified-site 466
 FT Region /note= "N-glycosylated"
 FT Region 540..658
 FT Region /label= HVR
 FT Region 659..880
 FT Region /label= H-3
 FT Region 667..688
 FT Region /label= U-3
 FT Misc-difference 672
 FT Region /note= "critical Asp residue"
 FT Modified-site 703
 FT Region /note= "N-glycosylated"
 FT Region 706..722
 FT Region /label= U-4
 FT Peptide 710..714
 FT Modified-site 846
 FT Region /note= "QXXRW motif"
 FT Region /note= "N-glycosylated"
 XX WO9818949-A2.
 XX 07-MAY-1998.
 XX 29-OCT-1997; 97WO-US019529.
 XX 29-OCT-1996; 96US-0029987P.
 XX (CALJ) CALGENE INC.
 XX Stalker DM, Pear JR, Delmer DP;
 XX WPI; 1998-272238/24.
 XX N-PSDB; AAV34432.
 XX New isolated plant cellulose synthesis enzyme DNA - is used to develop
 PT products for producing modified plant fibre products, particularly for
 PT modifying cotton fibre and wood quality.
 XX Example 5; Fig 3A-C; 49pp; English.
 XX This full-length polypeptide comprises cotton cellulose synthase CelA1.
 CC Its amino acid sequence was deduced from a cDNA clone (see AAV60039)
 CC isolated from a cotton fibre cDNA library. Expression of CelA1 and CelA2
 CC (see AAV60040) in developing fibre is approximately 500 times higher than
 CC the level of expression in other cotton tissues; approximately 1-2% of

CC the 24 dpa mRNA is Cels mRNA. Sequence analysis of CelA1 and CelA2 showed
 CC regions of high homology (H-1, H-2 and H-3) between plant and bacterial
 CC proteins, as well as a region (P-CR) that is conserved among plant Cels
 CC genes, and another region (HVR) that is hypervariable between plant
 CC genes. Regions U-1 to U-4 are predicted to be critical for UDP-glc
 CC binding and catalysis in bacterial Cels proteins. The isolated Cels cDNA
 CC clones are used in claimed methods for modifying fibre phenotype in a
 CC cotton plant and for modifying wood quality phenotype in a forest tree
 CC species. Lignin can be reduced by over-expressing the cellulose synthase
 CC product and diverting production into cellulose. (Updated on 17-OCT-2003
 CC to standardise OS field)
 XX SQ Sequence 974 AA;
 Query Match 84.7%; Score 50; DB 2; Length 974;
 Best Local Similarity 90.0%; Pred. No. 3e+02;
 Matches 9; Conservative 0; Mismatches 1; Indels 0; Gaps 0;
 QY 2 SSSSSSSCCC 11
 DB 553 SSSSSSSCCC 562
 ||||| |||||
 RESULT 7
 AAW73308
 ID AAW73308 standard; protein; 974 AA.
 XX AAW73308;
 XX 08-FEB-1999 (first entry)
 DT Cellulose synthase PcsA1.
 DE Cellulose synthase; cotton.
 KW Gossypium hirsutum.
 XX EP875575-A2.
 XX 04-NOV-1998.
 XX 31-MAR-1998; 98EP-00302489.
 XX 01-APR-1997; 97JP-00083133.
 XX (NISN) NISSHINBO IND INC.
 XX Hasegawa O, Aotsuka S, Ihara Y, Hayashi T;
 XX WPI; 1998-559440/48.
 XX N-PSDB; AAV08372.
 XX New DNA encoding cotton cellulose synthase proteins - useful for
 PT controlling cellulose synthesis in prokaryotic or eukaryotic cells.
 XX Claim 1; Page 14-18; 51pp; English.
 XX This sequence is a cellulose synthase of the invention. The DNA can be
 CC used for controlling cellulose synthesis in a cell by introducing the DNA
 CC into the cell, expressing RNA with a nucleotide sequence homologous to
 CC the DNA or a nucleotide sequence complementary to the DNA. The cells are
 CC prokaryotic cells i.e. acetobacterium and/or eukaryotic cells such as
 CC yeasts i.e. Saccharomyces spp., plant cells (cotton plants) and cultured
 CC cells of mammals etc
 XX SQ Sequence 974 AA;
 Query Match 84.7%; Score 50; DB 2; Length 974;
 Best Local Similarity 90.0%; Pred. No. 3e+02;
 Matches 9; Conservative 0; Mismatches 1; Indels 0; Gaps 0;
 QY 2 SSSSSSSCCC 11
 ||||| |||||

Db 553 SSSSSSCCCC 562

RESULT 8
ABO81687
ID ABO81687 standard; protein; 273 AA.
XX
XX ABO81687;
AC
XX
XX 29-JUL-2004 (first entry)
DT
XX
XX Pseudomonas aeruginosa polypeptide #13862.
DE
XX
XX Bacterial infection; Pseudomonas aeruginosa infection; antibacterial.
KW
XX
XX Pseudomonas aeruginosa.
OS
XX
XX US6551795-B1.
PN
XX
XX 22-APR-2003.
PD
XX
XX 18-FEB-1999; 99US-00252991.
PF
XX
XX 18-FEB-1998; 98US-0074788P.
PR
XX
XX 27-JUL-1998; 98US-0094190P.
PR
XX
XX (GENO-) GENOME THERAPEUTICS CORP.
PA
XX
XX Rubenfield MJ, Nolling J, Deloughery C, Bush D;
PI
XX
XX WPI; 2003-615309/58.
DR
XX
XX N-PSDB; ABD15258.
DR
XX
XX Novel isolated nucleic acid encoding Pseudomonas aeruginosa polypeptide,
PT
XX
XX useful as molecular targets for diagnostics, prophylaxis and treatment of
PT
XX
XX pathological conditions resulting from bacterial infection.
PT
XX
XX Disclosure; SEQ ID NO 30433; 455pp; English.
PS
XX
XX The invention relates to Pseudomonas aeruginosa polypeptides and the
CC
XX
XX polynucleotides encoding them. The sequences are useful in diagnosis and
CC
XX
XX therapy of pathological conditions, as molecular targets for diagnostics,
CC
XX
XX prophylaxis and treatment of pathological conditions resulting from a
CC
XX
XX bacterial infection, for evaluating a compound, such as a polypeptide,
CC
XX
XX for the ability to bind a P. aeruginosa nucleic acid, as components of
CC
XX
XX effective antibacterial targets, as targets for antibacterial drugs,
CC
XX
XX including anti-P. aeruginosa drugs, as templates for recombinant
CC
XX
XX production of P. aeruginosa-derived peptides or polypeptides, as target
CC
XX
XX components for diagnosis and/or treatment of P. aeruginosa-caused
CC
XX
XX infection, and in detection of P. aeruginosa sequences or other sequences
CC
XX
XX of Pseudomonas species using biochip technology. Sequences ABO67826-
CC
XX
XX ABO84396 represent P. aeruginosa polypeptides of the invention. Note: The
CC
XX
XX sequence data for this patent did not form part of the printed
CC
XX
XX specification but was obtained in electronic format from USPTO at
CC
XX
XX seqdata.uspto.gov/sequence.html
XX
XX
XX Sequence 273 AA;
SQ

Query Match 81.4%; Score 48; DB 7; Length 273;
Best Local Similarity 88.9%; Pred. No. 1.8e+02;
Matches 8; Conservative 1; Mismatches 0; Indels 0; Gaps 0;
QY 3 SSSSSSCCCC 11
Db 85 SSSSTCCCC 93

RESULT 9
ADK68328
ID ADK68328 standard; protein; 240 AA.
XX
XX
XX ADK68328;
AC
XX

Query Match 79.7%; Score 47; DB 7; Length 240;
Best Local Similarity 90.0%; Pred. No. 2.1e+02;
Matches 9; Conservative 1; Mismatches 0; Indels 0; Gaps 0;
QY 1 SSSSSSSCCC 10
Db 154 TSSSSSSCCC 163

RESULT 10
AD42828
ID AD42828 standard; protein; 240 AA.
XX
XX AD42828;
AC
XX
XX 22-APR-2004 (first entry)
DT
XX
XX Plant transcription factor #463.
DE
XX
XX transgenic; plant; enhanced tolerance to abiotic stress;
KW
XX
XX glyphosphate tolerance; hormone sensitivity; disease resistance;
KW
XX
XX sugar sensing; flowering; flower structure; stem bifurcation;
KW
XX
XX branching pattern; apical dominance; trichome; stem morphology;
KW
XX
XX root growth; root hair; seed development; cell proliferation;
KW
XX
XX cell differentiation; premature senescence; necrosis; plant size;
KW
XX
XX leaf morphology; seed morphology; seed biochemistry; root anthocyanin;
KW
XX
XX plant anthocyanin; light response; shade avoidance; bioinformatic;
KW
XX
XX transcription factor; gene; ds.
OS
XX
XX Atriplex hortensis.
XX
XX US2004019927-A1.
PN
XX
XX 29-JAN-2004.
PD
XX
XX 25-FEB-2003; 2003US-00374780.
PF
XX
XX

DT 06-MAY-2004 (first entry)
XX
XX Plant AP2/EREBP transcription factor SEQ ID NO:1.
DE
XX
XX plant; transcription factor.
KW
XX
XX Atriplex hortensis.
OS
XX
XX CN1400308-A.
PN
XX
XX 05-MAR-2003.
PD
XX
XX 01-AUG-2001; 2001CN-00123827.
PF
XX
XX 01-AUG-2001; 2001CN-00123827.
PR
XX
XX (GENE-) INST GENETICS CHINESE ACAD SCI.
PA
XX
XX Chen S, Liu Q, Shen Y;
PI
XX
XX WPI; 2003-422309/40.
DR
XX
XX N-PSDB; ADK68329.
DR
XX
XX Transcription factor from prunella asiatica, gene coding said factor and
PT
XX
XX method for cultivating plant with adverse resistance.
PT
XX
XX Claim 1; SEQ ID NO 1; 14pp; Chinese.
PS
XX
XX The invention relates to an AP2/EREBP transcription factor, the gene for
CC
XX
XX encoding AP2/EREBP, an expression vector containing the gene and the
CC
XX
XX method for breeding a resistant plant by using the gene and vector. The
CC
XX
XX present sequence represents the AP2/EREBP transcription factor of the
CC
XX
XX invention.
CC
XX
XX Sequence 240 AA;
SQ

```
PR 18-APR-2001; 2001US-00837944.
XX (SHER/) SHERMAN B K.
PA (RIEC/) RIECHMANN J L.
PA (JIAN/) JIANG C.
PA (HEAR/) HEARD J E.
PA (HAAK/) HAAKE V.
PA (CREE/) CREELMAN R A.
PA (RATC/) RATCLIFFE O.
PA (ADAM/) ADAM L J.
PA (REUB/) REUBER T L.
PA (KEDD/) KEDDIE J.
PA (BROU/) BROWN P E.
PA (PILG/) PILGRIM M L.
PA (DUBE/) DUBELL A N.
PA (PINE/) PINEDA O.
PA (YUGG/) YU G.
XX Sherman BK, Riechmann JL, Jiang C, Heard JE, Haake V;
PI Creelman RA, Ratcliffe O, Adam LJ, Reuber TL, Keddle J, Brown PE;
PI Pilgrim ML, Dubell AN, Pineda O, Yu G;
XX WPI; 2004-132245/13.
XX
XX New transgenic plant comprising a recombinant polynucleotide of any one
PT of more than 500 nucleotide sequences, useful in bioinformatic search
PT methods.
XX
XX Claim 1; SEQ ID NO 1291; 435pp; English.
XX
XX The invention describes a transgenic plant comprising a recombinant
CC polynucleotide of any one of more than 500 nucleotide sequences fully
CC defined in the specification or its complement. The method of the
CC invention can be used to produce a plant having altered traits such as:
CC enhanced tolerance to abiotic stress; glyphosphate tolerance; hormone
CC sensitivity; disease resistance; sugar sensing; early or late flowering;
CC altered flower structure, change in stem bifurcations, altered branching
CC pattern, reduced apical dominance, reduced trichome density; lack of
CC trichomes; reduced ectopic trichome development; altered trichome
CC development; increase in trichome number; altered stem morphology;
CC increased root growth; increased root hairs; altered seed development;
CC altered cell proliferation or cell differentiation; rapid development;
CC premature senescence; increased necrosis; increase in seedling or plant
CC size; decreased plant size; leaf morphology; seed morphology; seed
CC biochemistry; increase in root anthocyanins; increase in plant
CC anthocyanine, or alteration in light response or shade avoidance. The
CC transgenic plant, polynucleotides and polypeptides are useful in
CC bioinformatic search methods. This is the amino acid sequence of a plant
CC transcription factor, and an orthologue of Arabidopsis thaliana
CC transcription factors isolated in the invention, that can be used in the
CC creation of a transgenic plant with altered traits.
XX
XX Sequence 240 AA;
SQ
    Query Match          79.7%; Score 47; DB 8; Length 240;
    Best Local Similarity 90.0%; Pred. No. 2.1e+02;
    Matches 9; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

Qy 1 SSSSSSSSSCC 10
Db 154 TSSSSSSSSCC 163

RESULT 11
ADP30061
ID ADP30061 standard; protein; 298 AA.
XX
XX ADP30061;
XX
XX 12-AUG-2004 (first entry)
XX
XX Human secreted protein SEQ ID #828.
DE
XX
```

Cytostatic; Antiinflammatory; Immunosuppressive; Antibacterial; Virucide; cancer; inflammatory; immune; human secreted protein.

Homo sapiens.

WO2004035732-A2.

29-APR-2004.

28-AUG-2003; 2003WO-US026780.

29-AUG-2002; 2002US-0406576P.

29-AUG-2002; 2002US-0406579P.

29-AUG-2002; 2002US-0406585P.

29-AUG-2002; 2002US-0406588P.

29-AUG-2002; 2002US-0406608P.

29-AUG-2002; 2002US-0406611P.

29-AUG-2002; 2002US-0406612P.

29-AUG-2002; 2002US-0406616P.

29-AUG-2002; 2002US-0406640P.

29-AUG-2002; 2002US-0406642P.

29-AUG-2002; 2002US-0406646P.

29-AUG-2002; 2002US-0406653P.

29-AUG-2002; 2002US-0406655P.

29-AUG-2002; 2002US-0406666P.

17-SEP-2002; 2002US-0410946P.

17-SEP-2002; 2002US-0410947P.

17-SEP-2002; 2002US-0410948P.

17-SEP-2002; 2002US-0410949P.

17-SEP-2002; 2002US-0410953P.

17-SEP-2002; 2002US-0410957P.

17-SEP-2002; 2002US-0410958P.

17-SEP-2002; 2002US-0410959P.

17-SEP-2002; 2002US-0410960P.

17-SEP-2002; 2002US-0410961P.

17-SEP-2002; 2002US-0410962P.

17-SEP-2002; 2002US-0411019P.

17-SEP-2002; 2002US-0411022P.

17-SEP-2002; 2002US-0411023P.

17-SEP-2002; 2002US-0411024P.

17-SEP-2002; 2002US-0411032P.

17-SEP-2002; 2002US-0411035P.

17-SEP-2002; 2002US-0411037P.

17-SEP-2002; 2002US-0411041P.

17-SEP-2002; 2002US-0411045P.

17-SEP-2002; 2002US-0411046P.

17-SEP-2002; 2002US-0411048P.

17-SEP-2002; 2002US-0411052P.

17-SEP-2002; 2002US-0411055P.

17-SEP-2002; 2002US-0411073P.

17-SEP-2002; 2002US-0411082P.

17-SEP-2002; 2002US-0411101P.

17-SEP-2002; 2002US-0411111P.

18-APR-2003; 2003US-0463700P.

18-APR-2003; 2003US-0463708P.

18-APR-2003; 2003US-0463716P.

18-APR-2003; 2003US-0463732P.

02-MAY-2003; 2003US-0467199P.

02-MAY-2003; 2003US-0467201P.

02-MAY-2003; 2003US-0467203P.

02-MAY-2003; 2003US-0467230P.

19-MAY-2003; 2003US-0471306P.

19-MAY-2003; 2003US-0471336P.

22-MAY-2003; 2003US-0472420P.

22-MAY-2003; 2003US-0472430P.

09-JUN-2003; 2003US-0476609P.

09-JUN-2003; 2003US-0476641P.

08-JUL-2003; 2003US-0485218P.

08-JUL-2003; 2003US-0485223P.

08-JUL-2003; 2003US-0485224P.

08-JUL-2003; 2003US-0485252P.

14-JUL-2003; 2003US-0486446P.

14-JUL-2003; 2003US-0486480P.

PR 15-JUL-2003; 2003US-0486891P.
 PR 15-JUL-2003; 2003US-0486960P.
 PR 08-AUG-2003; 2003US-0493341P.
 PR 08-AUG-2003; 2003US-0493370P.
 PR 08-AUG-2003; 2003US-0493373P.
 PR 08-AUG-2003; 2003US-0493577P.
 PA (FIVE-) FIVE PRIME THERAPEUTICS INC.
 XX Williams LT, Chu K, Lee E, Hestir K, Beaurang PA, Behrens D;
 PI Halenbeck RF, Huang WM, Kothakota S, Haishan L, Linnemann T;
 PI Pierce K, Wang Y, Wong JGP, Wu G, Zhang H;
 XX WPI; 2004-348438/32.
 XX New nucleic acid molecule for diagnosing, preventing or treating diseases
 PT such as proliferative (e.g. cancer), inflammatory, immune, metabolic,
 PT genetic, bacterial and viral diseases.
 PS Claim 1; SEQ ID NO 2059; 428pp; English.
 XX The present invention relates to an isolated nucleic acid molecule
 CC encoding a polypeptide which is believed to be cytostatic,
 CC antiinflammatory, immunosuppressive, antibacterial and antiviral. The
 CC composition and methods are useful for diagnosing, preventing and
 CC treating diseases such as proliferative (e.g. cancer), inflammatory,
 CC immune, metabolic, genetic, bacterial and viral diseases. The present
 CC sequence represents a human secreted protein. The present sequence is
 CC available on WIPOWEB and is not in the specification.
 XX Sequence 298 AA;
 SQ
 Query Match 79.7%; Score 47; DB 8; Length 298;
 Best Local Similarity 100.0%; Pred. No. 2.6e+02;
 Matches 8; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
 QY 4 SSSSSCCC 11
 Db 43 SSSSSCCC 50
 |||||
 RESULT 12
 ABB92981
 ID ABB92981 standard; protein; 958 AA.
 XX AC ABB92981;
 XX DT 31-MAY-2002 (first entry)
 XX Herbicidally active polypeptide SEQ ID NO 2192.
 DE Herbicidal; plant; agriculture; herbicide.
 XX Arabidopsis thaliana.
 XX WO200210210-A2.
 XX 07-FEB-2002.
 XX 28-AUG-2001; 2001WO-EP009892.
 XX 28-AUG-2001; 2001WO-EP009892.
 XX (FARB) BAYER AG.
 XX Tietjen K, Weidler M;
 XX WPI; 2002-269010/31.
 XX Identifying plant target proteins for herbicidally active compounds,
 PT comprising aligning and comparing nucleic acid or amino acid sequences
 PT from plant with nucleic acid or amino acid sequences from non-plant
 PT organisms.

XX Claim 5; SEQ ID NO 2192; 261pp + Sequence Listing; English.
 XX The invention relates to identifying target proteins (ABB90790-ABB94016)
 CC for herbicidally active compounds, comprising aligning and comparing
 CC nucleic acid or amino acid sequences from plant with nucleic acid or
 CC amino acid sequences from non-plant organisms using suitable search
 CC parameters, where plant sequences having an E-value greater by a factor
 CC of 3 than the E-value of most similar non-plant sequences are selected.
 CC The polypeptides or nucleic acids encoding them are useful for
 CC identifying modulators. The identified modulators are useful as
 CC herbicides
 XX Sequence 958 AA;
 SQ
 Query Match 79.7%; Score 47; DB 5; Length 958;
 Best Local Similarity 100.0%; Pred. No. 6.6e+02;
 Matches 8; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
 QY 4 SSSSSCCC 11
 Db 555 SSSSSCCC 562
 |||||
 RESULT 13
 ADM06120
 ID ADM06120 standard; protein; 145 AA.
 XX AC ADM06120;
 XX DT 20-MAY-2004 (first entry)
 XX Human protein of the invention SEQ ID NO:4805.
 DE human; gene therapy; diagnostic marker; pharmaceutical.
 XX Homo sapiens.
 XX EP1347046-A1.
 XX 24-SEP-2003.
 XX 12-APR-2002; 2002EP-00008400.
 XX 22-MAR-2002; 2002JP-00137785.
 XX (REAS-) RES ASSOC BIOTECHNOLOGY.
 XX Isogai T, Sugiyama T, Otsuki T, Wakamatsu A, Sato H, Ishii S;
 PI Yamamoto J, Isono Y, Hio Y, Otsuka K, Nagai K, Irie R, Tamechika I;
 PI Seki N, Yoshikawa T, Otsuka M, Nagahari K, Masuho Y;
 XX WPI; 2003-723558/69.
 DR N-PSDB; ADM03677.
 XX New polynucleotides and polypeptides are useful in gene therapy, for
 PT developing a diagnostic marker or medicines for regulating their
 PT expression and activity, or as a target of gene therapy.
 XX Claim 1; SEQ ID NO 4805; 305pp; English.
 XX The invention relates to a novel human polynucleotide and the encoded
 CC polypeptide. A polynucleotide of the invention may have a use in gene
 CC therapy. An oligonucleotide of the invention ADM06202-ADM06773 is useful
 CC as a primer for synthesizing the polynucleotide or as a probe for
 CC detecting the polynucleotide. The polynucleotides ADM0316-ADM03758 are
 CC useful in gene therapy, for developing a diagnostic marker or medicines
 CC for regulating their expression and activity, or as a target of gene
 CC therapy. The proteins ADM03759-ADM06201 encoded by the polynucleotides
 CC are useful as pharmaceutical agents. The present sequence represents a
 CC protein sequence of the invention.
 XX Sequence 145 AA;
 SQ

Query Match 78.0%; Score 46; DB 7; Length 145;
 Best Local Similarity 72.7%; Pred. No. 1.9e+02;
 Matches 8; Conservative 1; Mismatches 2; Indels 0; Gaps 0;

Qy 1 SSSSSSSSSCCC 11
 || : |||||
 Db 50 SSPTSASSCCC 60

RESULT 14
 ABO80411
 ID ABO80411 standard; protein; 175 AA.
 XX AC ABO80411;
 XX DT 29-JUL-2004 (first entry)
 XX DE Pseudomonas aeruginosa polypeptide #12586.
 XX KW Bacterial infection; Pseudomonas aeruginosa infection; antibacterial.
 XX OS Pseudomonas aeruginosa.
 XX PN US6551795-B1.
 XX PD 22-APR-2003.
 XX PF 18-FEB-1999; 99US-00252991.
 XX PR 18-FEB-1998; 98US-0074788P.
 XX PR 27-JUL-1998; 98US-0094190P.
 XX PA (GENO-) GENOME THERAPEUTICS CORP.
 XX PI Rubenfield MJ, Nolling J, Deloughery C, Bush D;
 XX DR WPI; 2003-615309/58.
 XX DR N-PSDB; ABD13982.
 XX CC Novel isolated nucleic acid encoding Pseudomonas aeruginosa polypeptide,
 PT useful as molecular targets for diagnostics, prophylaxis and treatment of
 PT pathological conditions resulting from bacterial infection.
 XX PS Disclosure; SEQ ID NO 29157; 455pp; English.
 XX CC The invention relates to Pseudomonas aeruginosa polypeptides and the
 CC polynucleotides encoding them. The sequences are useful in diagnosis and
 CC therapy of pathological conditions, as molecular targets for diagnostics,
 CC prophylaxis and treatment of pathological conditions resulting from a
 CC bacterial infection, for evaluating a compound, such as a polypeptide,
 CC for the ability to bind a P. aeruginosa nucleic acid, as components of
 CC effective antibacterial targets, as targets for antibacterial drugs,
 CC including anti-P. aeruginosa drugs, as templates for recombinant
 CC production of P. aeruginosa-derived peptides or polypeptides, as target
 CC components for diagnosis and/or treatment of P. aeruginosa-caused
 CC infection, and in detection of P. aeruginosa sequences or other sequences
 CC of Pseudomonas species using biochip technology. Sequences ABO67826-
 CC ABO84396 represent P. aeruginosa polypeptides of the invention. Note: The
 CC sequence data for this patent did not form part of the printed
 CC specification but was obtained in electronic format from USPTO at
 CC seqdata.uspto.gov/sequence.html
 XX SQ Sequence 175 AA;

Query Match 78.0%; Score 46; DB 7; Length 175;
 Best Local Similarity 72.7%; Pred. No. 2.2e+02;
 Matches 8; Conservative 1; Mismatches 2; Indels 0; Gaps 0;

Qy 1 SSSSSSSSSCCC 11
 || : |||||
 Db 30 SPSPSASSCCC 40

RESULT 15
 AAP93703
 ID AAP93703 standard; protein; 216 AA.
 XX AC AAP93703;
 XX DT 25-MAR-2003 (revised)
 XX DT 29-OCT-1992 (first entry)
 XX DE Sequence of the 20kd surface antigen recognised by monoclonal antibody
 DE 6A5.
 XX KW Vaccine; coccidiosis; poultry; bivalent vaccine.
 XX OS Eimeria.
 XX PN BP344808-A.
 XX PD 06-DEC-1989.
 XX PF 02-JUN-1980; 80EP-00110056.
 XX PR 03-JUN-1988; 88US-00202721.
 XX PA (HOFF) HOFFMANN-LA ROCHE AG.
 XX PI Altenburge W, Binger MH, Chizzonite RA, Kramer RA, Lomedico PT;
 XX PI Mcandrew SJ;
 XX DR WPI; 1989-358220/49.
 XX DR N-PSDB; AAN92576.
 XX PT DNA sequences encoding Eimeria surface antigens - used in pox:virus
 PT vectors as a vaccine to protect chicks against coccidiosis.
 XX PS Claim 2; Fig 15; 78pp; English.
 XX CC The inventors claim a new protein which comprises one or more
 CC immunoreactive and/or antigenic determinants of an Eimeria surface
 CC antigen of mol. wt. 28,37,120 or more than 200 kD which specifically
 CC binds to one or more monoclonal antibody (Mab) from ATCC HB 9707-9712
 CC (see AAP93703-6). Also new are DNA encoding the protein (see AAN92576-9),
 CC and a vaccine comprising one or more proteins. Vaccine utility can be
 CC enhanced by inserting additional genes into the carrier virus (see
 CC AAP91652). "20 kd protein" means a recombinant or synthetic protein
 CC having an apparent mol. wt. of about 20 kilodaltons in SDS polyacrylamide
 CC gel electrophoresis which binds specifically to monoclonal antibody 6A5.
 CC This antibody also specifically reacts with an Eimeria surface antigen
 CC (from a whole extract of Eimeria proteins) having an apparent mol. wt. of
 CC about 28 kd in SDS gels. This antigen is present in the sporozoite
 CC development stage. The nucleotide sequence of a cDNA mol. encoding this
 CC protein and the amino acid sequence predicted therefrom are shown in
 CC AAN92576, AAN92577, AAP93703 and AAP93704. (Updated on 25-MAR-2003 to
 CC correct PF field.)
 XX SQ Sequence 216 AA;

Query Match 78.0%; Score 46; DB 1; Length 216;
 Best Local Similarity 100.0%; Pred. No. 2.6e+02;
 Matches 9; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 2 SSSSSSSSSCCC 10
 |||||
 Db 87 SSSSSSSSSCCC 95

Search completed: November 15, 2005, 12:22:55
 Job time : 89.1316 secs

GenCore version 5.1.6
Copyright (c) 1993 - 2005 CompuGen Ltd.

OM protein - protein search, using sw model

Run on: November 15, 2005, 12:11:54 ; Search time 9.8421 Seconds
(without alignments)
107.536 Million cell updates/sec

Title: US-10-820-998-4

Perfect score: 59

Sequence: 1 SSSSSSSSSCCC 11

Scoring table:

BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 283416 seqs, 96216763 residues

Total number of hits satisfying chosen parameters: 283416

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database :

PIR 79:.*
1: pir1:.*
2: pir2:.*
3: pir3:.*
4: pir4:.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	50	84.7	360	2 T45649	hypothetical prote
2	50	84.7	974	2 T10797	cellulose synthase
3	47	79.7	958	2 T04870	cellulose synthase
4	47	79.7	1671	2 S71828	sensory transducti
5	46	78.0	216	2 S05575	sporozoite antigen
6	46	78.0	313	2 T26465	hypothetical prote
7	45	76.3	1205	2 T41987	hypothetical prote
8	44	74.6	131	2 S50807	probable membrane
9	43	72.9	185	2 T05008	hypothetical prote
10	43	72.9	311	2 T45851	hypothetical prote
11	43	72.9	1193	2 A86193	hypothetical prote
12	42	71.2	220	2 D84420	hypothetical prote
13	42	71.2	385	1 OMRTSP	selenoprotein P pr
14	42	71.2	596	2 I38228	Sbb protein - huma
15	42	71.2	1023	2 JC4013	major acidic nucle
16	42	71.2	1193	2 G71605	hypothetical prote
17	42	71.2	1307	2 G96711	unknown protein, 9
18	41	69.5	170	2 S56958	probable membrane
19	41	69.5	185	2 T51003	hypothetical prote
20	41	69.5	198	2 A49243	GLORP-C4 protein -
21	41	69.5	230	2 D86352	protein T28f17.14
22	41	69.5	255	1 WJH04B	homeotic protein H
23	41	69.5	264	2 T30922	hypothetical prote
24	41	69.5	301	2 T01499	NADPH HC toxin red
25	41	69.5	333	1 JW0097	bicoid-related hom
26	41	69.5	369	2 I38559	melanoma antigen M
27	41	69.5	411	1 EDBEM4	45K immediate-earl
28	41	69.5	464	2 T51090	plastid division p
29	41	69.5	524	2 S33640	homeotic protein s

30 41 69.5 539 2 T15256
31 41 69.5 542 2 T06728
32 41 69.5 636 2 T06793
33 41 69.5 748 2 T49633
34 41 69.5 767 2 T05662
35 41 69.5 980 1 TVCTMD
36 40 67.8 135 2 A43098
37 40 67.8 260 2 B71353
38 40 67.8 268 2 T05971
39 40 67.8 275 2 T05969
40 40 67.8 325 2 T24689
41 40 67.8 380 2 T10442
42 40 67.8 427 2 S57776
43 40 67.8 444 2 T09474
44 40 67.8 455 2 T12041
45 40 67.8 583 1 S22544

hypothetical prote
pectate lyase (EC
receptor kinase ho
glucan 1,4-alpha-g
hypothetical prote
macrophage colony-
transcription fact
probable D,D-carbo
probable receptor-
receptor-like kina
hypothetical prote
selenoprotein P pr
cysteine proteinas
forkhead protein F
cysteine proteinas
transcription fact

ALIGNMENTS

RESULT 1

T45649

hypothetical protein F13112.180 - Arabidopsis thaliana

C;Species: Arabidopsis thaliana (mouse-ear cress)

C;Date: 04-Feb-2000 #sequence_revision 04-Feb-2000 #text_change 09-Jul-2004

C;Accession: T45649

R;Choisne, N.; Robert, C.; Brottier, P.; Wincker, P.; Artiguenave, F.; Sai

submitted to the Protein Sequence Database, November 1999

A;Reference number: Z23010

A;Accession: T45649

A;Status: preliminary

A;Molecule type: DNA

A;Residues: 1-360 <CHO>

A;Cross-references: UNIPROT:O9SD60; EMBL:AL133292

A;Experimental source: cultivar Columbia; BAC clone F13112

C;Genetics:

A;Map position: 3

A;Introns: 60/3

A;Note: F13112.180

Query Match 84.7%; Score 50; DB 2; Length 360;
Best Local Similarity 100.0%; Pred. No. 8.3;
Matches 10; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 SSSSSSSSSCC 10

Db 97 SSSSSSSSSCC 106

RESULT 2

T10797

cellulose synthase (EC 2.4.1.-) catalytic chain cel1 - upland cotton

C;Species: Gossypium hirsutum (upland cotton)

C;Date: 16-Jul-1999 #sequence_revision 16-Jul-1999 #text_change 09-Jul-2004

C;Accession: T10797

R;Pear, J.R.; Kawagoe, Y.; Schreckengost, W.E.; Delmer, D.P.; Stalker, D.M.

proc. Natl. Acad. Sci. U.S.A. 93, 12637-12642, 1996

A;Title: Higher plants contain homologs of the bacterial celA genes encoding the catalytic

A;Reference number: Z17152; MUID:97057296; PMID:8901635

A;Accession: T10797

A;Status: preliminary; translated from GB/EMBL/DBJ

A;Molecule type: mRNA

A;Residues: 1-974 <PEA>

A;Cross-references: UNIPROT:P93155; EMBL:U58283; NID:G1706955; PIDN:AAB37766.1; PID:G170

A;Experimental source: strain Acala SJ-2; fiber

C;Genetics:

A;Gene: cel1

C;Function:

A;Description: involved in the synthesis of cellulose

C;Keywords: glycosyltransferase; hexosyltransferase

Query Match 84.7%; Score 50; DB 2; Length 974;

Best Local Similarity 90.0%; Pred. No. 17;
Matches 9; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 2 SSSSSSSCCC 11
Db 553 SSSSSSSCCC 562
|||||

RESULT 3
T04870
cellulose synthase (EC 2.4.1.-) catalytic chain F28A21.190 - Arabidopsis thaliana
C;Species: Arabidopsis thaliana (mouse-ear cress)
C;Date: 23-Apr-1999 #sequence_revision 23-Apr-1999 #text_change 09-Jul-2004
C;Accession: T04870
R;Bavan, M.; Mueller, M.W.; Muendlein, A.; Felber, R.; Bancroft, I.; Mewes, H.W.; Mayer,
submitted to the Protein Sequence Database, February 1999
A;Reference number: Z15387
A;Accession: T04870
A;Molecule type: DNA
A;Residues: 1-958 <BRV>
A;Cross-references: UNIPROT:Q9SN37; EMBL:AL035526
A;Experimental source: cultivar Columbia; BAC clone F28A21
C;Genetics:
A;Map position: 4
A;Introns: 60/1; 76/3; 94/1; 149/2; 240/2; 355/3; 401/3; 443/3; 514/3; 662/3; 780/3; 791/3
A;Note: F28A21.190
C;Keywords: glycosyltransferase; hexosyltransferase

Query Match 79.7%; Score 47; DB 2; Length 958;
Best Local Similarity 100.0%; Pred. No. 39;
Matches 8; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 4 SSSSSSSCCC 11
Db 555 SSSSSSSCCC 562
|||||

RESULT 4
S71628
sensory transduction histidine kinase doka - slime mold (Dictyostelium discoideum)
C;Species: Dictyostelium discoideum
C;Date: 27-Nov-1997 #sequence_revision 12-Dec-1997 #text_change 09-Jul-2004
C;Accession: S71628; S78068
R;Schuster, S.C.; Noegel, A.A.; Oehme, F.; Gerisch, G.; Simon, M.I.
EMBO J 15, 3880-3889, 1996
A;Title: The hybrid histidine kinase Doka is part of the osmotic response system of Dictyostelium
A;Reference number: S71628; MUID:96324396; PMID:8670893
A;Accession: S71628
A;Status: nucleic acid sequence not shown
A;Molecule type: DNA
A;Residues: 1-1670 <SCH>
A;Cross-references: UNIPROT:Q23901; EMBL:X96869
R;Experimental source: strain AX2; substrain 214
R;Schuster, S.C.; Noegel, A.A.; Oehme, F.; Gerisch, G.; Simon, M.I.
submitted to the EMBL Data Library, March 1996
A;Description: The hybrid histidine kinase Doka is part of the osmotic response system of Dictyostelium
A;Reference number: S78068
A;Accession: S78068
A;Molecule type: DNA
A;Residues: 1-149, 'E', 151-219, 'TRVLKLIQSTNNIYWY', 238-1671 <SCW>
A;Cross-references: EMBL:X96869; NID:g1237201; PIDN:CAA65612.1; PID:g1237202
C;Genetics:
A;Gene: doka
C;Function:
A;Description: modulates cell response to changes in osmolarity; involved in spore formation
C;Keywords: phosphoprotein; signal transduction
F;1520-1629/Domain: response regulator homology <RRH2>
F;1568/Binding site: phosphate (Asp) (covalent) #status predicted

Query Match 79.7%; Score 47; DB 2; Length 1671;
Best Local Similarity 63.6%; Pred. No. 56;
Matches 7; Conservative 4; Mismatches 0; Indels 0; Gaps 0;

QY 1 SSSSSSSCCC 11
Db 655 NTNSSSSSSCCC 665
|||||

RESULT 5
S05575
sporozoite antigen - Eimeria tenella
C;Species: Eimeria tenella
C;Date: 07-Sep-1990 #sequence_revision 07-Sep-1990 #text_change 09-Jul-2004
C;Accession: S05575
R;Liberator, P.A.; Hsu, J.; Turner, M.J.
Nucleic Acids Res. 17, 7104, 1989
A;Title: Tandem trinucleotide repeats throughout the nucleotide sequence of a cDNA encoding
A;Reference number: S05575; MUID:89386062; PMID:2780319
A;Accession: S05575
A;Molecule type: mRNA
A;Residues: 1-216 <LIIB>
A;Cross-references: UNIPROT:P15744; EMBL:X15898

Query Match 78.0%; Score 46; DB 2; Length 216;
Best Local Similarity 100.0%; Pred. No. 18;
Matches 9; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 2 SSSSSSSCCC 10
Db 87 SSSSSSSCCC 95
|||||

RESULT 6
T26465
hypothetical protein Y11D7A.11 - Caenorhabditis elegans
C;Species: Caenorhabditis elegans
C;Date: 15-Oct-1999 #sequence_revision 15-Oct-1999 #text_change 09-Jul-2004
C;Accession: T26465
R;Steward, C.
submitted to the EMBL Data Library, October 1998
A;Reference number: Z20218
A;Accession: T26465
A;Status: preliminary; translated from GB/EMBL/DDBJ
A;Molecule type: DNA
A;Residues: 1-313 <WIL>
A;Cross-references: UNIPROT:Q9XWR2; EMBL:AL032632; PIDN:CAA21586.1; GSPDB:GN00022; CESP:Y11D7A
C;Genetics:
A;Gene: CESP:Y11D7A.11
A;Map position: 4
A;Introns: 26/2; 54/3; 88/1

Query Match 78.0%; Score 46; DB 2; Length 313;
Best Local Similarity 90.0%; Pred. No. 24;
Matches 9; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 1 SSSSSSSSSCCC 10
Db 110 SSSSSSSSDSCC 119
|||||

RESULT 7
T41987
hypothetical protein U86 - human herpesvirus 7 (strain JI)
C;Species: human herpesvirus 7
A;Variety: strain JI
C;Date: 03-Dec-1999 #sequence_revision 03-Dec-1999 #text_change 09-Jul-2004
C;Accession: T41987
R;Nicholas, J.
submitted to the EMBL Data Library, December 1995
A;Description: Determination and analysis of the complete nucleotide sequence of human herpesvirus 7
A;Reference number: Z22022
A;Accession: T41987
A;Status: preliminary; translated from GB/EMBL/DDBJ
A;Molecule type: DNA
A;Residues: 1-1205 <NIC>

A;Cross-references: UNIPROT:Q69513; EMBL:U43400; PIDN:AA54747.1
 A;Experimental source: strain J1
 C;Genetics:
 A;Note: U86

Query Match 76.3%; Score 45; DB 2; Length 1205;
 Best Local Similarity 81.8%; Pred. No. 80;
 Matches 9; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

QY 1 SSSSSSSSSCCC 11
 |||||
 Db 669 SSSSSSSSKCSC 679

RESULT 8

S50807
 probable membrane protein YJL064w - yeast (Saccharomyces cerevisiae)
 N;Alternate names: hypothetical protein HRC131; hypothetical protein J1120
 C;Species: Saccharomyces cerevisiae
 C;Date: 13-Jan-1995 #sequence_revision 08-Sep-1995 #text_change 09-Jul-2004
 A;Accession: S50807; S47126; S56838
 R;Vandenbol, M.; Durand, P.; Dion, C.; Portetelle, D.; Hilger, F.
 Yeast 11, 57-60, 1995

A;Title: Sequence of a 17.1 kb DNA fragment from chromosome X of Saccharomyces cerevisiae
 A;Reference number: S50798; MUID:95282514; PMID:7762302
 A;Accession: S50807
 A;Status: nucleic acid sequence not shown; translation not shown
 A;Molecule type: DNA
 A;Residues: 1-131 <NAV>

A;Cross-references: UNIPROT:P47038; EMBL:Z34288; NID:g498992; PIDN:CAA84058.1; PID:g49900
 A;Note: the nucleotide sequence was submitted to the EMBL Data Library, June 1994
 R;Vandenbol, M.; Durand, P.; Dion, C.; Portetelle, D.; Hilger, F.
 submitted to the EMBL Data Library, June 1994

A;Description: Sequence analysis of a 17.1 kb DNA fragment from chromosome X of Sacchar

A;Reference number: S47117

A;Accession: S47126

A;Molecule type: DNA

A;Residues: 1-131 <NAV>

A;Cross-references: EMBL:Z34288; NID:g498992; PID:g499002

R;Vandenbol, M.; Durand, P.; Portetelle, D.; Hilger, F.

submitted to the Protein Sequence Database, September 1995

A;Reference number: S56835

A;Accession: S56838

A;Molecule type: DNA

A;Residues: 1-131 <POH>

A;Cross-references: EMBL:Z49340; NID:g1008212; PID:g1008214; MIPS:YJL064w

C;Genetics:

A;Cross-references: SGD:S0003600

A;Map position: 10L

C;Superfamily: Saccharomyces probable membrane protein YJL064w

C;Keywords: transmembrane protein

Query Match 74.6%; Score 44; DB 2; Length 131;
 Best Local Similarity 63.8%; Pred. No. 23;
 Matches 7; Conservative 1; Mismatches 3; Indels 0; Gaps 0;

QY 1 SSSSSSSSSCCC 11
 |||||
 Db 49 SSSSGSGTCCC 59

RESULT 9

T05008
 hypothetical protein T19P19.100 - Arabidopsis thaliana
 C;Species: Arabidopsis thaliana (mouse-ear cress)
 C;Date: 23-Apr-1999 #sequence_revision 23-Apr-1999 #text_change 09-Jul-2004
 A;Accession: T05008

R;Bevan, M.; Monfort, A.; Casacuberta, E.; Puigdomenech, P.; Hoheisel, J.; Mewes, H.W.;

submitted to the Protein Sequence Database, April 1998

A;Reference number: Z15394

A;Accession: T05008

A;Molecule type: DNA

A;Residues: 1-185 <BEV>

A;Cross-references: UNIPROT:O65658; EMBL:AL022605
 A;Experimental source: cultivar Columbia; BAC clone T19P19

C;Genetics:

A;Map position: 4

A;Introns: 116/3; 156/1

A;Note: T19P19.100

Query Match 72.9%; Score 43; DB 2; Length 185;
 Best Local Similarity 81.8%; Pred. No. 39;
 Matches 9; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

QY 1 SSSSSSSSSCCC 11
 |||||
 Db 25 SSSSSSFSLC 35

RESULT 10

T45651
 hypothetical protein F13112.200 - Arabidopsis thaliana
 C;Species: Arabidopsis thaliana (mouse-ear cress)
 C;Date: 04-Feb-2000 #sequence_revision 04-Feb-2000 #text_change 09-Jul-2004
 A;Accession: T45651

R;Choisne, N.; Robert, C.; Bröttier, P.; Wincker, P.; Cattolico, L.; Artiguenave, F.; Sa

submitted to the Protein Sequence Database, November 1999

A;Reference number: Z23010

A;Accession: T45651

A;Status: preliminary

A;Molecule type: DNA

A;Residues: 1-311 <CHO>

A;Cross-references: UNIPROT:Q9SD58; EMBL:AL133292

A;Experimental source: cultivar Columbia; BAC clone F13112

C;Genetics:

A;Map position: 3

A;Introns: 160/2

A;Note: F13112.200

Query Match 72.9%; Score 43; DB 2; Length 311;
 Best Local Similarity 88.9%; Pred. No. 56;
 Matches 8; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

QY 2 SSSSSSSSSCCC 10
 |||||
 Db 87 SSSSTSSCC 95

RESULT 11

A86193
 hypothetical protein [imported] - Arabidopsis thaliana
 C;Species: Arabidopsis thaliana (mouse-ear cress)
 C;Date: 02-Mar-2001 #sequence_revision 02-Mar-2001 #text_change 09-Jul-2004
 A;Accession: A86193

R;Theologis, A.; Ecker, J.R.; Palm, C.J.; Federspiel, N.A.; Kaul, S.; White, O.; Alonso,

Chen, C.W.; Chung, M.K.; Conn, L.; Conway, A.B.; Conway, A.R.; Creasy, T.H.; Dewar, K.;

ansen, N.F.; Hughes, B.; Huizar, L.

Nature 408, 816-820, 2000

A;Authors: Hunter, J.L.; Jenkins, J.; Johnson-Hopson, C.; Khan, S.; Khaykin, E.; Kim, C.

C.A.; Li, J.H.; Li, X.; Liu, S.X.; Liu, Z.A.; Luros, J.S.; Maiti, R.; Marziani,

Rizzo, M.; Rooney, T.; Rowley, D.; Sakano, H.

A;Authors: Salzberg, S.L.; Schwartz, J.R.; Shinn, P.; Southwick, A.M.; Sun, H.; Tallon, I.

ker, M.; Wu, D.; Yu, G.; Fraser, C.M.; Venter, J.C.; Davis, R.W.

A;Title: Sequence and analysis of chromosome 1 of the plant Arabidopsis.

A;Reference number: A86141; MUID:21016719; PMID:11130712

A;Accession: A86193

A;Status: preliminary

A;Molecule type: DNA

A;Residues: 1-1193 <STO>

A;Cross-references: UNIPROT:Q9MA43; GB:AE005172; NID:g6850313; PIDN:AAF29390.1; GSPDB:GN

C;Genetics:

A;Map position: 1

Query Match 72.9%; Score 43; DB 2; Length 1193;
 Best Local Similarity 81.8%; Pred. No. 1.4e+02;
 Matches 9; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

GenCore version 5.1.6
Copyright (c) 1993 - 2005 Compugen Ltd.

OM protein - protein search, using sw model

Run on: November 15, 2005, 12:13:54 ; Search time 22.2895 Seconds
(without alignments)
36.840 Million cell updates/sec

Title: US-10-820-998-4
Perfect score: 59
Sequence: 1 SSSSSSSSSCCC 11

Scoring table: BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 513545 seqs, 74649064 residues

Total number of hits satisfying chosen parameters: 513545

Minimum DB seq length: 0
Maximum DB seq length: 2000000000
Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : Issued Patents AA: *
1: /cgn2_6/prodata1/iaa/5A COMB.pep.*
2: /cgn2_6/prodata1/iaa/5B COMB.pep.*
3: /cgn2_6/prodata1/iaa/6A COMB.pep.*
4: /cgn2_6/prodata1/iaa/6B COMB.pep.*
5: /cgn2_6/prodata1/iaa/PCUS COMB.pep.*
6: /cgn2_6/prodata1/iaa/backfiles1.pep.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	52	88.1	438	4	US-09-248-796A-23919
2	50	84.7	974	3	US-08-960-048-6
3	50	84.7	974	4	US-09-838-586-6
4	48	81.4	273	4	US-09-252-991A-30433
5	46	78.0	175	4	US-09-252-991A-29157
6	44	74.6	74	4	US-09-673-395A-168
7	44	74.6	131	4	US-09-538-092-443
8	44	74.6	143	4	US-09-252-991A-17703
9	44	74.6	327	4	US-09-252-991A-30618
10	44	74.6	847	4	US-10-162-012-2
11	43	72.9	262	4	US-09-252-991A-25990
12	43	72.9	323	4	US-09-252-991A-22747
13	42	71.2	218	1	US-08-463-115-92
14	42	71.2	218	1	US-08-463-388-92
15	42	71.2	219	1	US-08-463-115-91
16	42	71.2	219	1	US-08-463-388-91
17	42	71.2	264	1	US-08-463-115-93
18	42	71.2	264	1	US-08-463-388-93
19	42	71.2	331	4	US-09-252-991A-18772
20	42	71.2	384	4	US-09-252-991A-24427
21	42	71.2	594	4	US-09-949-016-9764
22	42	71.2	596	4	US-09-949-016-6303
23	41	69.5	29	3	US-09-136-769A-5
24	41	69.5	29	3	US-09-136-769A-16
25	41	69.5	131	3	US-08-981-392-60
26	41	69.5	131	4	US-09-908-322-60
27	41	69.5	133	4	US-09-252-991A-25367

28	41	69.5	148	4	US-09-640-211A-790	Sequence 790, Appl
29	41	69.5	156	4	US-09-252-991A-26582	Sequence 26582, A
30	41	69.5	177	4	US-09-270-767-46544	Sequence 46544, A
31	41	69.5	198	4	US-09-252-991A-19009	Sequence 19009, A
32	41	69.5	214	4	US-08-722-570-2	Sequence 2, Appli
33	41	69.5	214	4	US-08-932-411A-2	Sequence 2, Appli
34	41	69.5	249	4	US-09-248-796A-27128	Sequence 27128, A
35	41	69.5	255	4	US-09-949-016-6882	Sequence 6882, Ap
36	41	69.5	281	4	US-09-949-016-8978	Sequence 8978, Ap
37	41	69.5	301	3	US-09-231-227-4	Sequence 4, Appli
38	41	69.5	301	4	US-09-768-585-4	Sequence 34524, A
39	41	69.5	336	4	US-09-270-767-34524	Sequence 4, Appli
40	41	69.5	336	4	US-09-270-767-49741	Sequence 4, Appli
41	41	69.5	369	2	US-08-773-870-4	Sequence 6, Appli
42	41	69.5	391	1	US-08-602-010A-6	Sequence 6, Appli
43	41	69.5	391	1	US-08-680-726A-6	Sequence 6, Appli
44	41	69.5	391	3	US-09-092-409-6	Sequence 6, Appli
45	41	69.5	434	4	US-09-270-767-34959	Sequence 34959, A

ALIGNMENTS

RESULT 1

US-09-248-796A-23919
; Sequence 23919, Application US/09248796A
; Patent No. 6747137

; GENERAL INFORMATION:

; APPLICANT: Keith Weinstock et al

; TITLE OF INVENTION: NUCLEIC ACID AND AMINO ACID SEQUENCES RELATING TO CANDIDA ALBICANS
; TITLE OF INVENTION: FOR DIAGNOSTICS AND THERAPEUTICS

; FILE REFERENCE: 107196.132

; CURRENT APPLICATION NUMBER: US/09/248,796A

; PRIOR FILING DATE: 1999-02-12

; PRIOR APPLICATION NUMBER: US 60/074,725

; PRIOR FILING DATE: 1998-02-13

; PRIOR APPLICATION NUMBER: US 60/096,409

; PRIOR FILING DATE: 1998-08-13

; NUMBER OF SEQ ID NOS: 28208

; SEQ ID NO 23919

; LENGTH: 438

; TYPE: PRT

; ORGANISM: Candida albicans

; FEATURE:

; NAME/KEY: UNSURE

; LOCATION: (3),(4)

; OTHER INFORMATION: Identity of amino acid sequences at the above locations are unknown

US-09-248-796A-23919

Query Match 88.1%; Score 52; DB 4; Length 438;
Best Local Similarity 90.0%; Pred. No. 21;
Matches 9; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

QY 2 SSSSSSSSSCCC 11

Db 246 SASSSSSSSCCC 255

RESULT 2

US-08-960-048-6

; Sequence 6, Application US/08960048C

; Patent No. 6271443

; GENERAL INFORMATION:

; APPLICANT: Stalker, D. et al

; TITLE OF INVENTION: Plant Cellulose Synthase and Promoter

; TITLE OF INVENTION: Sequences

; FILE REFERENCE: 15621/01/US

; CURRENT APPLICATION NUMBER: US/08/960,048C

; CURRENT FILING DATE: 1997-10-29

; PRIOR APPLICATION NUMBER: 60/029,987

; PRIOR FILING DATE: 1996-10-29

; NUMBER OF SEQ ID NOS: 12

; SOFTWARE: FastSeq for Windows Version 4.0

; SEQ ID NO 6
; LENGTH: 974
; TYPE: PRT
; ORGANISM: Gossypim hirsutum
US-08-960-048-6

Query Match 84.7%; Score 50; DB 3; Length 974;
Best Local Similarity 90.0%; Pred. No. 74;
Matches 9; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 2 SSSSSSSCCC 11
|||||:||||
DB 553 SSSSSSCCCC 562

RESULT 3

US-09-838-586-6

; Sequence 6, Application US/09838586

; Patent No. 6576818

; GENERAL INFORMATION:

; APPLICANT: Stalker, D. et al.

; TITLE OF INVENTION: Plant Cellulose Synthase and Promoter

; FILE REFERENCE: 15621/02/US

; CURRENT APPLICATION NUMBER: US/09/838,586

; PRIOR FILING DATE: 2001-04-18

; PRIOR FILING DATE: 1996-10-29

; PRIOR FILING DATE: 1997-10-29

; NUMBER OF SEQ ID NOS: 12

; SOFTWARE: Fast-SEQ for Windows Version 4.0

; SEQ ID NO 6

; LENGTH: 974

; TYPE: PRT

; ORGANISM: Gossypim hirsutum

US-09-838-586-6

Query Match 84.7%; Score 50; DB 4; Length 974;
Best Local Similarity 90.0%; Pred. No. 74;
Matches 9; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 2 SSSSSSSCCC 11
|||||:||||
DB 553 SSSSSSCCCC 562

RESULT 4

US-09-252-991A-30433

; Sequence 30433, Application US/09252991A

; Patent No. 6551795

; GENERAL INFORMATION:

; APPLICANT: Marc J. Rubenfield et al.

; TITLE OF INVENTION: NUCLEIC ACID AND AMINO ACID SEQUENCES RELATING TO PSEUDOMONAS

; FILE REFERENCE: 107196.136

; CURRENT APPLICATION NUMBER: US/09/252,991A

; PRIOR FILING DATE: 1999-02-18

; PRIOR FILING DATE: 1998-02-18

; PRIOR FILING DATE: 1998-02-18

; PRIOR FILING DATE: 1998-07-27

; NUMBER OF SEQ ID NOS: 33142

; SEQ ID NO 30433

; LENGTH: 273

; TYPE: PRT

; ORGANISM: Pseudomonas aeruginosa

US-09-252-991A-30433

Query Match 81.4%; Score 48; DB 4; Length 273;
Best Local Similarity 88.9%; Pred. No. 42;
Matches 8; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

QY 3 SSSSSSSCCC 11
|||||:||||
DB 85 SSSSSSCCCC 93

RESULT 5

US-09-252-991A-29157

; Sequence 29157, Application US/09252991A

; Patent No. 6551795

; GENERAL INFORMATION:

; APPLICANT: Marc J. Rubenfield et al.

; TITLE OF INVENTION: NUCLEIC ACID AND AMINO ACID SEQUENCES RELATING TO PSEUDOMONAS

; FILE REFERENCE: 107196.136

; CURRENT APPLICATION NUMBER: US/09/252,991A

; PRIOR FILING DATE: 1999-02-18

; PRIOR FILING DATE: 1998-02-18

; PRIOR FILING DATE: 1998-02-18

; PRIOR FILING DATE: 1998-07-27

; NUMBER OF SEQ ID NOS: 33142

; SEQ ID NO 29157

; LENGTH: 175

; TYPE: PRT

; ORGANISM: Pseudomonas aeruginosa

US-09-252-991A-29157

Query Match 78.0%; Score 46; DB 4; Length 175;
Best Local Similarity 72.7%; Pred. No. 48;
Matches 8; Conservative 1; Mismatches 2; Indels 0; Gaps 0;

QY 1 SSSSSSSSCCCC 11
|||||:||||
DB 30 SPSPSASSCCC 40

RESULT 6

US-09-673-395A-168

; Sequence 168, Application US/09673395A

; Patent No. 6620923

; GENERAL INFORMATION:

; APPLICANT: SPECHT, THOMAS

; APPLICANT: HINZMANN, BERND

; APPLICANT: SCHMITT, ARMIN

; APPLICANT: PILARSKY, CHRISTIAN

; APPLICANT: DAHL, EDGAR

; APPLICANT: ROSENTHAL, ANDRE

; TITLE OF INVENTION: HUMAN NUCLEIC ACID SEQUENCES FROM UTERUS TUMOR TISSUE

; FILE REFERENCE: ALBRE-12

; CURRENT APPLICATION NUMBER: US/09/673,395A

; CURRENT FILING DATE: 2000-10-17

; NUMBER OF SEQ ID NOS: 637

; SOFTWARE: PatentIn Ver. 2.1

; SEQ ID NO 168

; LENGTH: 74

; TYPE: PRT

; ORGANISM: Homo sapiens

US-09-673-395A-168

Query Match 74.8%; Score 44; DB 4; Length 74;
Best Local Similarity 80.0%; Pred. No. 39;
Matches 8; Conservative 2; Mismatches 0; Indels 0; Gaps 0;

QY 1 SSSSSSSSCCCC 10
|||||:||||
DB 31 SSSASTSSCC 40

RESULT 7

US-09-538-092-443

; Sequence 443, Application US/09538092

; Patent No. 6753314

; GENERAL INFORMATION:


```

; APPLICANT: Giot, Loic
; APPLICANT: Mansfield, Traci A.
; TITLE OF INVENTION: Protein-Protein Complexes and Method of Using Same
; FILE REFERENCE: 15966-542
; CURRENT APPLICATION NUMBER: US/09/538,092
; CURRENT FILING DATE: 2000-03-29
; PRIOR APPLICATION NUMBER: 60/127,352
; PRIOR FILING DATE: 1999-04-01
; PRIOR APPLICATION NUMBER: 60/178,965
; PRIOR FILING DATE: 2000-02-01
; NUMBER OF SEQ ID NOS: 1387
; SOFTWARE: CurapatSeqFormatter Version 0.9
; SEQ ID NO 443
; LENGTH: 131
; TYPE: PRT
; ORGANISM: Saccharomyces cerevisiae
; FEATURE:
; NAME/KEY: misc feature
; LOCATION: (0)...(0)
; OTHER INFORMATION: Polypeptide Accession Number YJL064W
US-09-538-092-443

Query Match      74.6%; Score 44; DB 4; Length 131;
Best Local Similarity 63.6%; Pred. No. 64;
Matches      7; Conservative 1; Mismatches 3; Indels 0; Gaps 0;

Qy      1 SSSSSSSSSCCC 11
      ||||| : ||||
Db      49 SSSGSGGTCCC 59

RESULT 8
US-09-252-991A-17703
; Sequence 17703, Application US/09252991A
; Patent No. 6551795
; GENERAL INFORMATION:
; APPLICANT: Marc J. Rubenfield et al.
; TITLE OF INVENTION: NUCLEIC ACID AND AMINO ACID SEQUENCES RELATING TO PSEUDOMONAS
; TITLE OF INVENTION: AERUGINOSA FOR DIAGNOSTICS AND THERAPEUTICS
; FILE REFERENCE: 107196.136
; CURRENT APPLICATION NUMBER: US/09/252,991A
; CURRENT FILING DATE: 1999-02-18
; PRIOR APPLICATION NUMBER: US 60/074,788
; PRIOR FILING DATE: 1998-02-18
; PRIOR APPLICATION NUMBER: US 60/094,190
; PRIOR FILING DATE: 1998-07-27
; NUMBER OF SEQ ID NOS: 33142
; SEQ ID NO 17703
; LENGTH: 143
; TYPE: PRT
; ORGANISM: Pseudomonas aeruginosa
US-09-252-991A-17703

Query Match      74.6%; Score 44; DB 4; Length 143;
Best Local Similarity 70.0%; Pred. No. 69;
Matches      7; Conservative 2; Mismatches 1; Indels 0; Gaps 0;

Qy      2 SSSSSSSSSCCC 11
      : |||| : ||||
Db      41 ASSSRASCCC 50

RESULT 9
US-09-252-991A-30618
; Sequence 30618, Application US/09252991A
; Patent No. 6551795
; GENERAL INFORMATION:
; APPLICANT: Marc J. Rubenfield et al.
; TITLE OF INVENTION: NUCLEIC ACID AND AMINO ACID SEQUENCES RELATING TO PSEUDOMONAS
; TITLE OF INVENTION: AERUGINOSA FOR DIAGNOSTICS AND THERAPEUTICS
; FILE REFERENCE: 107196.136
; CURRENT APPLICATION NUMBER: US/09/252,991A
; CURRENT FILING DATE: 1999-02-18

```

; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 2
; LENGTH: 847
; TYPE: PRT
; ORGANISM: Homo sapiens
US-10-162-012-2

Query Match 74.6%; Score 44; DB 4; Length 847;
Best Local Similarity 72.7%; Pred. No. 3.3e+02;
Matches 8; Conservative 0; Mismatches 3; Indels 0; Gaps 0;

Qy 1 SSSSSSSSSCC 11
Db 160 SSPLSGSSCC 170

RESULT 11

US-09-252-991A-25990
; Sequence 25990, Application US/09252991A
; Patent No. 6551795
; GENERAL INFORMATION:
; APPLICANT: Marc J. Rubenfield et al.
; TITLE OF INVENTION: NUCLEIC ACID AND AMINO ACID SEQUENCES RELATING TO PSEUDOMONAS
; FILE REFERENCE: 107196.136
; CURRENT APPLICATION NUMBER: US/09/252,991A
; PRIOR FILING DATE: 1999-02-18
; PRIOR APPLICATION NUMBER: US 60/074,788
; PRIOR FILING DATE: 1998-02-18
; PRIOR APPLICATION NUMBER: US 60/094,190
; PRIOR FILING DATE: 1998-07-27
; NUMBER OF SEQ ID NOS: 33142
; SEQ ID NO 25990
; LENGTH: 262
; TYPE: PRT
; ORGANISM: Pseudomonas aeruginosa
US-09-252-991A-25990

Query Match 72.9%; Score 43; DB 4; Length 262;
Best Local Similarity 70.0%; Pred. No. 1.6e+02;
Matches 7; Conservative 1; Mismatches 2; Indels 0; Gaps 0;

Qy 2 SSSSSSSSSCC 11
Db 237 SPSASSGCC 246

RESULT 12

US-09-252-991A-22747
; Sequence 22747, Application US/09252991A
; Patent No. 6551795
; GENERAL INFORMATION:
; APPLICANT: Marc J. Rubenfield et al.
; TITLE OF INVENTION: NUCLEIC ACID AND AMINO ACID SEQUENCES RELATING TO PSEUDOMONAS
; FILE REFERENCE: 107196.136
; CURRENT APPLICATION NUMBER: US/09/252,991A
; PRIOR FILING DATE: 1999-02-18
; PRIOR APPLICATION NUMBER: US 60/074,788
; PRIOR FILING DATE: 1998-02-18
; PRIOR APPLICATION NUMBER: US 60/094,190
; PRIOR FILING DATE: 1998-07-27
; NUMBER OF SEQ ID NOS: 33142
; SEQ ID NO 22747
; LENGTH: 323
; TYPE: PRT
; ORGANISM: Pseudomonas aeruginosa
US-09-252-991A-22747

Query Match 72.9%; Score 43; DB 4; Length 323;
Best Local Similarity 100.0%; Pred. No. 1.9e+02;
Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 5 SSSSSCC 11
Db 93 SSSSSCC 99

RESULT 13

US-08-463-115-92
; Sequence 92, Application US/08463115
; Patent No. 5703221
; GENERAL INFORMATION:
; APPLICANT: WILLIAM JOHN MARTIN
; TITLE OF INVENTION: ISOLATED STEALTH VIRUSES
; NUMBER OF SEQUENCES: 104
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Lyon & Lyon
; STREET: 633 West Fifth Street
; CITY: Los Angeles
; STATE: California
; COUNTRY: U.S.A.
; ZIP: 90071-2066
; COMPUTER READABLE FORM:
; MEDIUM TYPE: 3.5" Diskette, 1.44 Mb
; MEDIUM TYPE: storage
; OPERATING SYSTEM: IBM P.C. DOS 5.0
; SOFTWARE: FastSeq Version 1.5
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/463,115
; FILING DATE: June 5, 1995
; CLASSIFICATION: 435
; PRIOR APPLICATION DATA:
; PRIOR APPLICATION DATA: including application
; PRIOR APPLICATION DATA: described below:
; APPLICATION NUMBER: 08/157,811
; FILING DATE: NO. 5703221ember 23, 1993
; APPLICATION NUMBER: 07/887,502
; FILING DATE: May 22, 1992
; APPLICATION NUMBER: 07/704,814
; FILING DATE: May 23, 1991
; APPLICATION NUMBER: 07/763,039
; FILING DATE: September 20, 1991
; ATTORNEY/AGENT INFORMATION:
; NAME: Warburg, Richard J.
; REGISTRATION NUMBER: 32,327
; REFERENCE/DOCKET NUMBER: 213/301
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (213) 489-1600
; TELEFAX: (213) 955-0440
; TELEX: 67-3510
; INFORMATION FOR SEQ ID NO: 92:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 218 amino acids
; TYPE: amino acid
; STRANDEDNESS: single
; TOPOLOGY: linear
US-08-463-115-92

Query Match 71.2%; Score 42; DB 1; Length 218;
Best Local Similarity 80.0%; Pred. No. 1.7e+02;
Matches 8; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

Qy 1 SSSSSSSSSCC 10
Db 169 SSSSSSECC 178

RESULT 14

US-08-465-388-92
; Sequence 92, Application US/08465388
; Patent No. 5753488
; GENERAL INFORMATION:

APPLICANT: WILLIAM JOHN MARTIN
TITLE OF INVENTION: ISOLATED STEALTH VIRUSES
NUMBER OF SEQUENCES: 104
CORRESPONDENCE ADDRESS:
ADDRESSEE: Lyon & Lyon
STREET: 633 West Fifth Street
STREET: Suite 4700
CITY: Los Angeles
STATE: California
COUNTRY: U.S.A.
ZIP: 90071-2066
COMPUTER READABLE FORM:
MEDIUM TYPE: 3.5" Diskette, 1.44 Mb
COMPUTER: IBM Compatible
OPERATING SYSTEM: IBM P.C. DOS 5.0
SOFTWARE: FastSeq Version 1.5
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/465,388
FILING DATE: June 5, 1995
CLASSIFICATION: 435
PRIOR APPLICATION DATA: including application
PRIOR APPLICATION DATA: described below: Four
APPLICATION NUMBER: 08/157,811
FILING DATE: NO. 5753488ember 23, 1993
APPLICATION NUMBER: 07/887,502
FILING DATE: May 22, 1992
APPLICATION NUMBER: 07/704,814
FILING DATE: May 23, 1991
APPLICATION NUMBER: 07/763,039
FILING DATE: September 20, 1991
ATTORNEY/AGENT INFORMATION:
NAME: Warburg, Richard J.
REGISTRATION NUMBER: 32,327
REFERENCE/DOCKET NUMBER: 213/300
TELECOMMUNICATION INFORMATION:
TELEPHONE: (213) 489-1600
TELEFAX: (213) 955-0440
TELEX: 67-3510
INFORMATION FOR SEQ ID NO: 92:
SEQUENCE CHARACTERISTICS:
LENGTH: 218 amino acids
TYPE: amino acid
STRANDEDNESS: single
TOPOLOGY: linear
US-08-465-388-92

Query Match 71.2%; Score 42; DB 1; Length 218;
Best Local Similarity 80.0%; Pred. No. 1.7e+02;
Matches 8; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

Qy 1 SSSSSSSSSCC 10
Db 169 SSSSSSECC 178

RESULT 15

US-08-463-115-91
Sequence 91, Application US/08463115
Patent No. 5703221
GENERAL INFORMATION:
APPLICANT: WILLIAM JOHN MARTIN
TITLE OF INVENTION: ISOLATED STEALTH VIRUSES
NUMBER OF SEQUENCES: 104
CORRESPONDENCE ADDRESS:
ADDRESSEE: Lyon & Lyon
STREET: 633 West Fifth Street
STREET: Suite 4700
CITY: Los Angeles
STATE: California

COUNTRY: U.S.A.
ZIP: 90071-2066
COMPUTER READABLE FORM:
MEDIUM TYPE: 3.5" Diskette, 1.44 Mb
COMPUTER: IBM Compatible
OPERATING SYSTEM: IBM P.C. DOS 5.0
SOFTWARE: FastSeq Version 1.5
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/463,115
FILING DATE: June 5, 1995
CLASSIFICATION: 435
PRIOR APPLICATION DATA: including application
PRIOR APPLICATION DATA: described below: Four
APPLICATION NUMBER: 08/157,811
FILING DATE: NO. 5703221ember 23, 1993
APPLICATION NUMBER: 07/887,502
FILING DATE: May 22, 1992
APPLICATION NUMBER: 07/704,814
FILING DATE: May 23, 1991
APPLICATION NUMBER: 07/763,039
FILING DATE: September 20, 1991
ATTORNEY/AGENT INFORMATION:
NAME: Warburg, Richard J.
REGISTRATION NUMBER: 32,327
REFERENCE/DOCKET NUMBER: 213/301
TELECOMMUNICATION INFORMATION:
TELEPHONE: (213) 489-1600
TELEFAX: (213) 955-0440
TELEX: 67-3510
INFORMATION FOR SEQ ID NO: 91:
SEQUENCE CHARACTERISTICS:
LENGTH: 219 amino acids
TYPE: amino acid
STRANDEDNESS: single
TOPOLOGY: linear
US-08-463-115-91

Query Match 71.2%; Score 42; DB 1; Length 219;
Best Local Similarity 80.0%; Pred. No. 1.7e+02;
Matches 8; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

Qy 1 SSSSSSSSSCC 10
Db 170 SSSSSSECC 179

Search completed: November 15, 2005, 12:29:36
Job time : 23.2895 secs

This Page Blank (uspto)